DOCUMENT RESUME

ED 288 892 TM 870 625

AUTHOR	Johnson, Eugene G.; And Others
TITLE	National Assessment of Educational Progress: 1985-86 Public-Use Data Tapes, Version 1.0. Users' Guide.
INSTITUTION	National Assessment of Educational Progress, Princeton, NJ.
SPONS AGENCY	Center for Education Statistics (OERI/ED), Washington, DC.
PUB DATE	Aug 87
GRANT	NIE-G-83-0011
NOTE	140p.; For related "Codebooks and Layouts," see TM 870 626-628.
AVAILABLE FROM	National Assessment of Educational Progress, Educational Testing Service, Rosedale Road, CN 6710, Princeton, NJ 08541-6710 ($25.00).
PUB TYPE	Guides - Non-Classroom Use (055)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	Academic Achievement; Computer Software; *Computer Storage Devices; Data Analysis; *Databases; *Data Collection; Data Interpretation; Educational Assessment; Elementary Secondary Education; Latent Trait Theory; *Magnetic Tapes; *National Competency Tests; National Surveys; Questionnaires; Reading Tests; Sampling; *Statistical Analysis; Testing Programs; Writing Evaluation
IDENTIFIERS	Balanced Incomplete Block Spiralling; *National Assessment of Educational Progress

ABSTRACT

This document is the users' guide for Version 1.0 of the Public-Use data tapes compiled by the National Assessment of Educational Progress (NAEP), 1985-86. The Public-Use tapes are produced to allow outside researchers access to the NAEP data. The tapes accompanying this guide, one for grade 3/age 9, one for grade 7/age 13, and one for grade 11/age 17, contain student responses to mathematics, science, computer competence, U.S. history, and literature exercises, as well as responses to questionnaires from students, teachers, and principals. Reading scores have been withheld from this 1.0 version, but will be included in later versions. The guide includes: (1) background information on NAEP and the public-use tapes; (2) special considerations for users; (3) description of the assessment instrument design, sample selection, data collection, and scoring procedures; (4) reporting subgroups and other variables; (5) suggestions on conducting statistical analyses of the NAEP data; (6) content and format of the data tapes; and (7) suggestions for using the Statistical Package for the Social Sciences-X (SPSS-X) and SAS computer software systems with this data. Appendices include a brief NAEP history and lists of related machine-readable data files and printed reports. (JGL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED288892

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS
1985-86 PUBLIC-USE DATA TAPES
VERSION 1.0
USERS' GUIDE

---

by

Eugene G. Johnson
Debra L. Kline
Norma A. Norris
Alfred M. Rogers

with contributions by

Laurel Barnett
John L. Barone
John J. Ferris
David S. Freund
Bruce A. Kaplan
Benjamin F. King
Janet R. Johnson
Robert J. Mislevy
Rebecca Zwick

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E. Driscoll

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

August 1987

Prepared for the
Office for Educational Research and Improvement,
Center for Education Statistics
under Grant No. NIE-G-83-0011
by the National Assessment of Educational Progress
Educational Testing Service, Princeton, New Jersey

TM 870 625

BEST COPY AVAILABLE

NOTICE

The 1985-86 National Assessment of Educational Progress included an assessment of reading achievement at three grade/age levels. Because of changes in the assessment procedures, however, the psychometric properties of the reading results are not fully understood at present. Accordingly, reading achievement results have not been included in the Version 1.0 data tapes pending the results of currently ongoing analyses of the properties of the reading data.

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS
1985-86 PUBLIC-USE DATA TAPES
VERSION 1.0
USERS' GUIDE

---

by

Eugene G. Johnson
Debra L. Kline
Norma A. Norris
Alfred M. Rogers

with contributions by

Laurel Barnett
John L. Barone
John J. Ferris
David S. Freund
Bruce A. Kaplan
Benjamin F. King
Janet R. Johnson
Robert J. Mislevy
Rebecca Zwick

August 1987

Prepared for the
Office for Educational Research and Improvement,
Center for Education Statistics
under Grant No. NIE-G-83-0011
by the National Assessment of Educational Progress
Educational Testing Service, Princeton, New Jersey

The National Assessment of Educational Progress is funded by the U. S. Department of Education under a grant to Educational Testing Service. National Assessment is an education research project mandated by Congress to collect data over time on the performance of young Americans in various learning areas. It makes available information on assessment procedures to state and local education agencies.

The Foundations of Literacy Project is funded by the National Endowment for the Humanities under a grant to the Educational Excellence Network, Columbia University. The Educational Excellence Network subcontracted the assessment of 17-year-old and 11th grade students' knowledge of history and literature to NAEP.

This document is based upon work performed pursuant to Grant No. NIE-G-83-0011 of the Office for Educational Research and Improvement, Center for Education Statistics. It does not, however, necessarily reflect the views of that agency.

NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS
1985-86 PUBLIC-USE DATA TAPES, VERSION 1.0

USERS' GUIDE

Table of Contents

Index of Tables ix

Chapter 1 INTRODUCTION 3

1.1 What is NAEP? 3
1.2 The 1985-86 NAEP 3
  1.2.1 Sponsorship 3
  1.2.2 Special Features 4
1.3 The NAEP Public-Use Data Tapes 4
1.4 The Package 4
1.5 Security and Confidentiality 6
1.6 Inquiries and Tape Orders 6
1.7 How to Use the Guide 7

Chapter 2 SPECIAL CONSIDERATIONS FOR USERS 11

2.1 Introduction 11
2.2 Grade/Age Sampling 11
2.3 Spiral Sample 11
2.4 Bridge Samples 11
2.5 Reporting Subgroups and Other Variables 12
2.6 Using Weights 12
2.7 Error Estimation 12

Chapter 3 INSTRUMENT DESIGN 15

3.1 Introduction 15
3.2 Assessment Items 15
  3.2.1 Mathematics 16
  3.2.2 Science 16
  3.2.3 Reading 17
  3.2.4 Computer Competence 17
  3.2.5 Literature 17
  3.2.6 U.S. History 18
  3.2.7 Common Background 18
3.3 Assembling Items into Instruments 19
  3.3.1 Assembling Items into Blocks 20
  3.3.2 Assembling Blocks into Booklets 20
  3.3.3 Spiralling and Bundling Booklets 21

3.4 Bridge Instruments (Booklets 1-5) 22
  3.4.1 Change in Age Definition and Time of Testing (Bridge A: Booklets 1-3) 23
  3.4.2 Change in Mode of Administration (Bridge B: Booklets 4-5) 23
3.5 Questionnaires 24
  3.5.1 Teacher 24
  3.5.2 School Characteristics and Policy 24
  3.5.3 Excluded Student 25
  3.5.4 Computer Coordinator 25
3.6 Tabular Summary of Assessment Instruments 25

Chapter 4 SAMPLE SELECTION 43

4.1 Introduction 43
4.2 Grade/Age Definition 43
4.3 Bridge Samples 44
4.4 Sample Design 44
  4.4.1 Primary Sampling Units 45
  4.4.2 Schools 45
  4.4.3 Assigning Sessions to Schools 46
  4.4.4 Sampling Students 47
  4.4.5 Sampling Excluded Students 47
  4.4.6 Sampling Teachers 47
  4.4.7 Sampling School Characteristics and Computer Coordinators 48
4.5 Sample Weights 48
  4.5.1 Base Weight 49
  4.5.2 Adjustment of Base Weights for Nonresponse 49
  4.5.3 Trimming of Weights 51
  4.5.4 Post-Stratification 52
  4.5.5 The Final Student-Weight: The Full Sample Weight 52
  4.5.6 Weights Used for Variance Estimation 53
4.6 Tabular Summary of the 1985-86 Samples 53

Chapter 5 DATA COLLECTION AND SCORING 63

5.1 Introduction 63
5.2 Field Administration and Data Collection 63
5.3 Data Entry System 63
  5.3.1 Editing 64
  5.3.2 Multiple Responses 64
5.4 Professionally Scored Items 65
5.5 Quality Control 65

Chapter 6 REPORTING SUBGROUPS AND OTHER VARIABLES 69

6.1 Introduction 69
6.2 Reporting Subgroups 69
  6.2.1 Sex 69
  6.2.2 Observed Race/Ethnicity 69
  6.2.3 Imputed Race/Ethnicity 70
  6.2.4 Size and Type of Community 71
  6.2.5 Parental Education 73
  6.2.6 Region 74
  6.2.7 Grade 75
  6.2.8 Age 75
6.3 Other Variables 77
  6.3.1 Principal Questionnaire 77
  6.3.2 Quality Education Data (QED) 78

Chapter 7 CONDUCTING STATISTICAL ANALYSES OF 1985-86 NAEP DATA 81

7.1 Introduction 81
7.2 Using Weights to Account for Differential Representation 82
  7.2.1 The Spiral Samples of Students 83
  7.2.2 The Bridge B Samples of Students 84
  7.2.3 The Excluded Students from the Spiral and Bridge B Assessments 84
  7.2.4 The Bridge A Samples of Students 85
  7.2.5 The Excluded Students from the Bridge A Assessments 85
  7.2.6 Other Weights 85
7.3 Procedures Used by NAEP to Estimate Variability (Jackknifing) 85
  7.3.1 Degrees of Freedom of the Jackknifed Variance Estimate 88
7.4 Approximations 89
7.5 A Note Concerning Multiple Comparisons 91

Chapter 8 CONTENT AND FORMAT OF THE DATA TAPES 95

8.1 Introduction 95
8.2 Raw Data Files 95
8.3 Layouts 95
8.4 Codebooks 98
8.5 Machine-Readable Catalogs 100
8.6 SPSS-X and SAS Control Statement Files 103

Chapter 9 WORKING WITH SPSS-X AND SAS 109

9.1 Introduction 109
9.2 SPSS-X and SAS Control Statement Files 109
9.3 Creating SPSS-X System Files 110
9.4 Creating SAS System Files 112
9.5 Merging Files Under SPSS-X or SAS 114
9.6 Computing the Estimated Variance of a Mean (Jackknifing) Using SPSS-X or SAS 114

Appendix A NAEP HISTORY 123

Appendix B RELATED MACHINE-READABLE DATA FILES AND PRINTED REPORTS 129

Index of Tables

Table 3-1 Potential and achieved block pairings across learning areas 21
3-2 Block information, grade 3/age 9 26
3-3 Block information, grade 7/age 13 27
3-4 Block information, grade 11/age 17 28
3-5 Booklet content, grade 3/age 9 30
3-6 Booklet content, grade 7/age 13 31
3-7 Booklet content, grade 11/age 17 32
3-8 Target sample sizes by grade/age 33
3-9 Comparison matrix, grade 3/age 9 spiral booklets 34
3-10 Comparison matrix, grade 7/age 13 spiral booklets 35
3-11 Comparison matrix, grade 11/age 17 spiral booklets 36
3-12 Number of students administered each spiral block and bridge booklet, grade 3/age 9 37
3-13 Number of students administered each spiral block and bridge booklet, grade 7/age 13 38
3-14 Number of students administered each spiral block and bridge booklet, grade 11/age 17 39

Table 4-1 The sampling subuniverses 45
4-2 Number of students by grade/age and assessment type 56
4-3 Number of spiral students, grade 3/age 9 57
4-4 Number of spiral students, grade 7/age 13 58
4-5 Number of spiral students, grade 11/age 17 59

Table 6-1 Determining imputed race/ethnicity 72
6-2 Geographic regions 74
6-3 Comparison of MODGRD and DGRADE values 76
6-4 Comparison of MODAGE and DAGE values 76

Table 8-1 Code definitions 96
8-2 Block designations and file locations 99
8-3 Machine-readable catalog data layout 101
8-4 Data file description, grade 3/age 9 104
8-5 Data file description, grade 7/age 13 105
8-6 Data file description, grade 11/age 17 106

Table 9-1 SPSS-X control statement synopsis 111
9-2 SAS control statement synopsis 113
9-3 Matching school and student files 116
9-4 Matching teacher and student files 116
9-5 Matching school and excluded student files 117
9-6 Matching school, teacher and student files 118
9-7 SPSS-X multi-weight standard error computation 119
9-8 SAS multi-weight standard error computation 120

Table A-1 NAEP learning areas, grades, and ages assessed, 1969-1986 125

Chapter 1

INTRODUCTION

Chapter 1: INTRODUCTION

1.1 WHAT IS NAEP?

The National Assessment of Educational Progress (NAEP) is a continuing, congressionally mandated national survey of the knowledge, skills, understandings, and attitudes of young Americans in major learning areas usually taught in school. Its primary goals are to detect and report the current status of, as well as changes in, the educational attainments of young Americans, and to report long-term trends in those attainments. The purpose of NAEP is to gather information which will aid educators, legislators, and others in improving the educational experience of youth in the United States. It is the first ongoing effort to obtain comprehensive and dependable achievement data on a national basis in a uniform, scientific manner.

NAEP began in 1969 to survey American students aged 9, 13, and 17 annually in various subject areas; young adults aged 26 to 35 were surveyed less frequently. Since the 1980-81 school year, budget restraints have prompted a shift to biennial data collection. In 1983-84, NAEP began sampling students by grade as well as age. More information about NAEP and its history are included in Appendix A.

1.2 THE 1985-86 NAEP

The 1985-86 assessment is a sample survey of achievement and attitudes in six learning areas: mathematics, science, reading, computer competence, literature, and U.S. history. National samples were drawn of students aged 9, 13, and 17 and of the corresponding modal grades 3, 7, and 11. Achievement in literature and U.S. history was assessed for grade 11/age 17 students only.

The data were collected between November 4, 1985 and May 2, 1986; the majority of data collection activity occurred between February 17 and May 2, 1986.

1.2.1 Sponsorship

The assessment was conducted by Educational Testing Service for the Office for Educational Research and Improvement, Center for Education Statistics. Assessments of achievement and attitudes in mathematics, science, reading, and computer understanding were funded by the U.S. Department of Education. The survey of achievement and attitudes of grade 11/age 17 students in literature and U.S. history was subcontracted to NAEP by the Educational Excellence Network under a grant from the National Endowment for the Humanities.

1.2.2 Special Features

Because of the complexity of the NAEP design (see Chapters 3 and 4), data tape users should have some understanding of the design before performing analyses. For example, because 1985-86 NAEP data were collected by both age and grade, the user must decide which sample is appropriate for a particular analysis. Special characteristics of the assessment are outlined in Chapter 2.

The data tapes contain sampling weights for each student. These weights should be used in statistical analyses. In addition, because of the complex sampling scheme, conventional methods of standard error estimation do not produce optimum estimates. The NAEP sampling scheme also reduces the effective degrees of freedom for statistical analysis. These issues are discussed in Chapter 7.

1.3 THE NAEP PUBLIC-USE DATA TAPES

Since 1975, NAEP has provided assessment data on computer tapes for use by outside researchers. In 1983 the format of the data tapes was refined to make them easier to use. Version 1.0 of the 1985-86 data tapes can be used for descriptive and item level analyses. To enhance secondary analyses of mathematics, reading, and science data, plausible values, which estimate individual proficiencies, will be added to Version 2.0 of the data tapes, scheduled for release in 1988.

These data tapes contain data only for the student and school samples and weights only for the student samples; data for excluded students and teachers associated with students, and weights for schools, excluded students, and teachers associated with students will be included in the second version. The Version 2.0 data tapes will also include scores for open-ended computer competence items.

1.4 THE PACKAGE

In addition to the computer tapes, NAEP provides a users' guide, printed layouts and codebooks describing each data file, microfiche copies of item text, data file printouts, and learning area objectives booklets.

Data Tapes

Three data tapes, one for each grade/age, are included in the package. Version 1.0 of the data tapes contains:

- student responses to mathematics, science, computer competence, U.S. history, and literature exercises (Note: Achievement results for reading have been withheld from the Version 1.0 tapes--see the notice at the front of this guide);

- responses to questions about students' backgrounds, attitudes, and activities;
- information about students' schools;
- sampling weights for each student; and
- SPSS-X and SAS control statement files.

The data tapes are available with the following standard tape characteristics:

Recording Density:	1600 or 6250 bytes per inch
Recording Format:	EBCDIC or ASCII
Blocking:	Blocked or unblocked
Label:	IBM standard, unlabeled, or ANSI

Content and format of the tapes are described in Chapter 8.

Codebooks

For each tape a codebook is included which provides the layout of the data, a description of each variable, and a description of each raw data file. The codebooks and their use are discussed in Chapter 8.

Printouts

Computer printouts containing the first 50 records of each data file are also included in the package.

Microfiche

In addition to the response data contained on the tapes, the text of all assessment items and survey questionnaires are provided on microfiche. These are color-coded by grade/age.

Objectives Booklets

NAEP develops educational objectives in all the areas it assesses. The booklets containing the objectives for the learning areas assessed in 1985-86 are included in the package.

1.5 SECURITY AND CONFIDENTIALITY

NAEP releases some items for unrestricted public use. Other items are kept secure so they can be re-administered in future assessments to permit analysis of trends in performance levels. To preserve the integrity of NAEP, it is essential that these items remain secure. For the 1985-86 assessment, the complete contents of the survey questionnaires will be released, as will all of the background and attitude items administered to students. However, achievement items in all learning areas will be held secure.

It is crucial to the long-term utility of NAEP that secure items not be published or used in other assessments or research projects. At the same time, the utility of the data for secondary researchers is seriously compromised if item text is unavailable. Therefore, microfiche copies of the complete exercises have been included. To protect the confidentiality of secure items, users are asked to sign a nondisclosure agreement when ordering the tape package.

A second confidentiality issue involves subregional identifiers. To prevent identification of individual respondents (which would violate the Privacy Act), all subregional identifiers (schools, districts, counties, states, etc.) have been deleted from data files. Scrambled primary sampling unit (PSU), school, and respondent identification fields are included on the data files to permit unique identification of each record and to permit aggregating data across exercise booklets at the school or PSU level.

1.6 INQUIRIES AND TAPE ORDERS

A list of data tapes available for previous NAEP assessments is provided in Appendix B. If you have questions about the tapes and their use, or want to order a tape package, contact one of the following individuals.

Subject-Matter Inquiries

Douglas Rhodes
Associate Director, NAEP
Educational Testing Service
Rosedale Road
Princeton, NJ 08541
(609) 734-1464

Computer-Related Inquiries and Tape Orders

Norma Norris
Senior Research Data Analyst
Educational Testing Service
Rosedale Road
Princeton, NJ 08541
(609) 734-5898

Tapes may be ordered for all grade/age groups or for individual grade/age groups. Orders will be shipped with machine-readable data files, a users' guide, printed layouts and codebooks, microfiche copies of item texts, objectives booklets, and a printout of the first 50 records for each data file.

1.7 HOW TO USE THE GUIDE

Chapters 2 through 9 and the appendices provide detailed information about the 1985-86 assessment, the data tapes, and recommended methods of working with the data to perform analyses. A summary of these chapters follows.

Chapter 2: SPECIAL CONSIDERATIONS FOR USERS

The design features of the assessment that may be of special concern to researchers who wish to perform their own analyses of the data.

Chapter 3: INSTRUMENT DESIGN

The six learning areas assessed; assessment item descriptions; the contents of survey questionnaires; methods by which items were administered to students; and a tabular summary of the results of item administration.

Chapter 4: SAMPLE SELECTION

The methods by which schools, students and teachers were chosen to be included in the assessment; the method by which some students were chosen for the sample but subsequently excluded from the assessment; the sampling weights included on the data tapes; and a tabular summary of the results of the sample selection.

Chapter 5: DATA COLLECTION AND SCORING

Procedures used for administering assessments; methods used for data entry and editing; how items were professionally scored; and methods used for quality control of the data.

Chapter 6: REPORTING SUBGROUPS AND OTHER VARIABLES

A description of the variables used for reporting; how reporting variables were derived; and a discussion of other variables on the tapes that are not self-explanatory.

Chapter 7: CONDUCTING STATISTICAL ANALYSES WITH NAEP DATA

A discussion of the weights on the data tapes and how to use them in different types of analyses; and methods for estimating standard errors and variances.

Chapter 8: CONTENT AND FORMAT OF THE DATA TAPES

A detailed description of the raw data files, layouts, codebooks, machine-readable catalogs, SPSS-X and SAS control statement files, printouts, and microfiche.

Chapter 9: WORKING WITH SPSS-X AND SAS

Procedures for creating SPSS-X and SAS system files, merging files, and using the jackknife procedure to estimate standard errors.

APPENDIX A provides information about the history of NAEP.

APPENDIX B lists machine-readable data files available for previous assessments and printed NAEP reports available for mathematics, reading, and science.

Chapter 2

SPECIAL CONSIDERATIONS FOR USERS

Chapter 2: SPECIAL CONSIDERATIONS FOR USERS

2.1 Introduction

Because of the complexity of the NAEP design, it is important for users to have some understanding of the design before performing analyses of the data. The following sections highlight areas of potential importance to the user in constructing analyses.

2.2 Grade/Age Sampling

The 1985-86 main assessment sampled students by grade as well as age, and includes three student cohorts: students who were either in the third grade or nine years old (grade 3/age 9), students who were either in the seventh grade or thirteen years old (grade 7/age 13), and students who were either in the eleventh grade or seventeen years old (grade 11/age 17).

Each weighted sample of students of a given grade/age, when combined with the appropriate weighted sample of students excluded from participation, is a representative sample of students of that grade/age in the national population. In most cases, the grade/age estimate is of minor interest; a researcher is more likely interested in the number of students at either a particular grade or a particular age.

Chapter 6 provides information about the NAEP variables that are used to determine grade and age for reporting purposes; Chapter 4 discusses grade/age definition and the resulting student samples.

2.3 Spiral Sample

The term "spiral" is a short name for the Balanced Incomplete Block (BIB) Spiral method of assembling assessment items into instruments. This method was developed to allow the study of the interrelationships among all items within a learning area. As a result of this design, even though all items are given to approximately the same number of students, no student receives all items.

The spiral design is discussed in Chapter 3, Section 3.3.

2.4 Bridge Samples

In addition to the spiral, or main, sample, the 1985-86 assessment also included a number of bridge samples which were used to assess the effect of certain changes from past assessments: changing the age definition of the two younger student cohorts, changing the time of year the two younger student cohorts were tested, and changing the mode of administration for all three cohorts from tape recorder-assisted to pencil-and-paper.

The instruments used for the bridge assessments are discussed in Chapter 3; the bridge samples are discussed in Chapters 4 and 7.

2.5 Reporting Subgroups and Other Variables

In addition to reporting national results, NAEP reports information by several student subgroups. Some subgroup data (for example, student ethnicity) are derived from responses to two or more assessment items. Chapter 6 defines the reporting subgroups and explains how their data are derived.

Some variables on the tapes are not taken from assessment instruments but from other sources. For optimal use of the data from these variables, please see their explanations in Chapter 6.

2.6 Using Weights

In the NAEP sampling design, students do not have an equal probability of being selected. Therefore, as in all complex surveys, each student has been assigned a sampling weight. When computing descriptive statistics or conducting inferential procedures, one should properly weight the data for each student. Performance of statistical analyses without weights can lead to misleading results.

The weighted number of students responding to the spiral assessment or to any bridge booklet, in combination with the appropriate corresponding samples of students excluded from the assessment, is an estimate of the number of students in the national population.

Chapter 7 provides an explanation of the weights on the data tapes and how to use them in performing analyses.

2.7 Error Estimation

The NAEP sampling design involves the selection of clusters of students who come from the same school, as well as clusters of schools that come from the same geographical region. As a result, observations are not independent of one another as they are in a simple random sample. Therefore, use of ordinary formulas for estimating the standard error of sample statistics will result in values that are too small. Alternate methods of computing standard errors are provided in Chapter 7.

Another effect of the sampling design is a reduction of the effective degrees of freedom, which in NAEP are a function of the number of clusters of primary sampling units and the number of strata in the design, rather than the number of subjects. Recommended formulas for obtaining degrees of freedom can be found in Chapter 7.

Chapter 3

INSTRUMENT DESIGN

Chapter 3: INSTRUMENT DESIGN[1]

3.1 INTRODUCTION

In 1985-86, NAEP conducted an assessment of achievement and attitudes of over 100,000 grade 3/age 9, grade 7/age 13, and grade 11/age 17 students in the learning areas of reading, mathematics, science, and computer competence, and over 10,000 grade 11/age 17 students in literature and U.S. history.

The assessment incorporated five types of instruments: student assessment booklets, a teacher questionnaire, a school characteristics and policies questionnaire, a computer coordinator questionnaire, and a questionnaire for excluded students.

This chapter will answer the following questions about the instruments used in the 1985-86 assessment:

- What content areas were assessed in each learning area? How were attitudes and interests assessed? What is a common background item? (3.2)
- How were items arranged into assessment booklets? What is BIB spiralling? (3.3)
- What are "bridge" assessment instruments and why were they administered? (3.4)
- What are "excluded students"? What instruments were used to collect data for teachers, schools, computer coordinators, and excluded students? (3.5)

For information about how schools, students, and teachers were selected for participation in the assessment, definition of the grade and age categories included in the assessment, and final counts of participants, see Chapter 4, "Sample Selection."

3.2 ASSESSMENT ITEMS

Of the six 1985-86 learning areas, mathematics, science, reading, and literature were assessed in earlier years; assessments of computer competence and U.S. history were conducted by NAEP for the first time in 1985-86.

---

[1] Data collected from the Teacher Questionnaire and the Excluded Student Questionnaire, and sampling weights for schools, excluded students, and teachers associated with students are not contained on the Version 1.0 data tapes, but will be included in Version 2.0.

Items were created from a set of objectives developed for each learning area. For mathematics, science, and reading, some items used in previous assessments were used again in 1985-86 to permit analyses of trends over time.

In each learning area, items were developed both to assess academic achievement (achievement items) and investigate student attitudes, experiences, and interests (attitude items). (NAEP booklets describing item development objectives for each learning area are included in the public-use data tapes package.)

In addition, a common core of items was administered to each student to collect data about the student's personal and family background.

The following sections summarize the content of assessment items for each learning area and the core of common background items.

3.2.1 Mathematics

Earlier assessments in mathematics were conducted during the school terms ending in 1973, 1978, and 1982. The objectives for development of mathematics items for 1985-86 were organized into seven content areas:

- Mathematical methods
- Discrete mathematics
- Data organization and interpretation
- Measurement
- Geometry
- Relations, functions, and algebraic expressions
- Numbers and operations

Some mathematics items required the use of a calculator, which was provided.

Five categories of attitude items were also developed:

- Mathematics in school
- Mathematics and oneself
- Mathematics and society
- Mathematics as a discipline
- Attitudes toward computers

3.2.2 Science

Earlier assessments in science were conducted during the school terms ending in 1970, 1973, 1977, and 1982. The objectives for development of science items for 1985-86 were organized into six content areas:

- Life sciences
- Physics
- Chemistry
- Earth and space sciences
- History of science
- Nature of science

Seven categories of attitude items were also developed:

- Attitudes toward science classes
- Career and education intentions
- Socioscientific responsibility
- Science as a personal tool
- Value of science
- Societal issues
- Experiences in science

3.2.3 Reading

Earlier assessments in reading were conducted during the school terms ending in 1971, 1975, 1980, and 1984. The objectives for development of reading items for 1985-86 were:

- Comprehends what is read
- Extends comprehension
- Manages the reading experience
- Values reading

The reading attitude items investigated what students read, both in and out of school; how often they read different kinds of material; how often they read for enjoyment; use of the library; understanding the value of reading; and the reading behavior of people in the students' homes.

3.2.4 Computer Competence

Computer competence was assessed by NAEP for the first time in 1985-86. Students were assessed at the cognitive levels of knowledge, operation, and problem solving/design in three content areas: knowledge and attitudes, programming, and computer applications. The eight applications areas are:

- Word processing
- Database management
- Lab instrumentation
- Telecommunications
- Graphics
- Music generation
- Spreadsheets
- Models and simulations

Students' attitudes were assessed on subjects such as computers' importance in society and in the workplace; computing opportunities in school; relevance of computer instruction to future success; and use of computers outside the school curriculum.

3.2.5 Literature

Grade 11/age 17 students' knowledge of literature was assessed by NAEP in 1985-86 as part of the Foundations of Literacy Project, funded by the

National Endowment for the Humanities under a grant to the Educational Excellence Network, Columbia University. Although NAEP conducted previous assessments of literature in 1970-71 and 1979-80, none of the items used in those assessments are common to the items used in 1985-86.

Achievement items were developed to assess knowledge within four literary genres:

- Novels, short stories, and plays
- Myths, epics, and biblical characters and stories
- Poetry
- Nonfiction

Students were also asked to respond to items about their experiences with literature in school and outside of school.

3.2.6 U.S. History

Grade 11/age 17 students' knowledge of U.S. history was assessed by NAEP for the first time in 1985-86, as part of the Foundations of Literacy Project, funded by the National Endowment for the Humanities under a grant to the Educational Excellence Network, Columbia University. Achievement items were developed to assess knowledge of six historical time periods:

- Exploration and colonization to 1763
- The Revolutionary War and the New Republic, 1763-1815
- Nationhood, Sectionalism, and the Civil War, 1815-1877
- Territorial expansion, the rise of modern America, and World War I, 1877-1920
- The Great Depression, the New Deal, and World War II, 1920-1945
- Post-World War II, 1945 to present

Students were also asked to respond to items about their school studies and activities in various areas of U.S. history.

3.2.7 Common Background

In addition to learning area achievement and attitude items, each student was asked to respond to general background items concerning subjects such as materials in the home, languages spoken, levels of parents' education, hours spent watching television, and after-school activities.

At grade 3/age 9 the common background questions were read aloud to students by the exercise administrator; this took fifteen minutes of assessment time. At the other two grade/ages, students were given six minutes to complete the common background questions. Only the first question, regarding student ethnicity, was read aloud. Students read and completed the remaining questions themselves.

The set of common background items differed slightly for each grade/age, but within each grade/age the same set of background items was used for every student.

3.3 ASSEMBLING ITEMS INTO ASSESSMENT INSTRUMENTS

In conducting its first NAEP assessment in 1983-84, Educational Testing Service incorporated significant changes in the design of the assessment instruments (see A New Design for a New Era by Messick, Beaton, and Lord, 1983). The new design, with some modifications, was used again in the 1985-86 assessment and is described in the following paragraphs.

The approach used prior to 1983-84 divided the entire pool of items designated for a given age group into a number of distinct sets, each of which would take a student about three-quarters of an hour to complete. Since no student was administered more than one set of items, this simple matrix design allowed the calculation of measure of relation among items within the same set but not among items in different sets.

To allow the study of the interrelationships among all items within a learning area, ETS assigned items to students by means of a Balanced Incomplete Block (BIB) design with spiralled administration. Items within a learning area were assembled into sixteen-minute blocks (thirteen minutes for grade 3/age 9), each block comprising approximately two minutes of attitude items and approximately fourteen minutes (eleven minutes for grade 3/age 9) of achievement items. Each student was administered a booklet containing three learning area blocks and a block of common background items, for a total testing time of approximately 54 minutes.

The balanced part of this method assigns blocks to booklets in such a way that each learning area block appears in the same number of booklets and each pair of blocks within a learning area appears in at least one booklet. Such a design generates a large number of different booklets.

The incomplete part of the method refers to the fact that no individual receives all items, thus yielding incomplete data for the respondents.

The spiralling part of the method cycles the order of booklets for administration so that typically no two students in any assessment session in a school, and at most only a few students in schools with multiple sessions, receive the same booklet. Spiralling thus reduces school clustering effects (see Chapter 4) and results in more efficient sampling.

To permit the calculation of correlations among items between learning areas, a design was formulated to provide for systematic pairing of blocks between learning areas, allowing selected items in each area to be linked to items in each of the other areas.

3.3.1 Assembling Items into Blocks

The first step in assembling assessment instruments was the grouping of items from each learning area into units called "blocks".

Blocks were created from the item pool for each learning area within each grade/age. An item was selected to be placed in a block based on the time required to complete the item. More blocks were assigned for math and science at successive grade/age levels because of the increasing range of coursework in those areas as grades advance.

Each resulting learning area block contained approximately fourteen minutes of achievement items (eleven minutes for grade 3/age 9) and two minutes of attitude items. Each block was given a unique identification number consisting of cohort age (9, 13, or 17), learning area initial (M for mathematics, R for reading, etc.), and sequence number (1 through 11).

Common background items were grouped into one block for each grade/age.

The resulting total of blocks and corresponding block identification numbers for each grade/age are presented in Tables 3-2, 3-3, and 3-4 at the end of the chapter. Six reading blocks, two mathematics blocks, and two science blocks were common to both grade 7/age 13 and grade 11/age 17.

3.3.2 Assembling Blocks into Booklets

The second step in assembling assessment instruments was the grouping of learning area blocks and common background blocks into booklets.

This assembly resulted in 51 different booklets for grade 3/age 9; 67 different booklets for grade 7/age 13; and 92 different booklets for grade 11/age 17. Each booklet for a given grade/age contained three learning area blocks and the appropriate common background block.

The learning area blocks used in each booklet for the three grade/ages are shown in Tables 3-5, 3-6, and 3-7 at the end of the chapter.

Methods of booklet assembly for each learning area are described in the following paragraphs.

Mathematics, Science, Reading, and Computer Competence

Blocks for these learning areas were assigned to booklets in such a way that, for a given grade/age, each learning area block appeared in the same number of booklets and each pair of blocks within a learning area appeared in at least one booklet.

Systematic and judicious pairing of certain blocks of items between learning areas was performed to permit selected items in each area to be linked to items in the other three areas. For example, at grade 3/age 9, where there were six reading blocks and seven mathematics blocks, the total possible block pairs between those learning areas was 42. Fifteen of those

27

pairings were achieved through the booklet design. Table 3-1 presents the achieved pairings of blocks across all learning areas for each grade/age.

Table 3-1

Potential and Achieved Block Pairings Across Learning Areas by Grade/Age

	GRADE 3/AGE 9 PAIRINGS		GRADE 7/AGE 13 PAIRINGS		GRADE 11/AGE 17 PAIRINGS	
LEARNING AREAS	Potential	Achieved	Potential	Achieved	Potential	Achieved
Reading-Mathematics	42	15	54	11	66	12
Reading-Science	42	15	54	11	66	14
Reading-Computer	18	12	36	20	36	28
Mathematics-Science	49	18	81	18	121	28
Mathematics-Computer	21	9	54	11	66	14
Science-Computer	21	9	54	11	66	12

Literature and U.S. History

Four special booklets were created for the administration of literature and U.S. history items to grade 11/age 17 students. Each of these booklets (numbered 92 through 95) contained one of the four literature blocks; one of the four U.S. history blocks; reading block 13R$_4$; and the grade 11/age 17 common background block. Reading block 13R$_4$, which is also used in other booklets, was included to permit correlations between reading ability and literature and history achievement.

3.3.3 Spiralling and Bundling Assessment Booklets

The order of the booklets for each grade/age was spiralled, or cycled for administration in such a way that, typically, no two students in any one assessment session received the same booklet. Booklets were packaged together in bundles of 29. (Note: Booklets 1 through 5 were not included in the bundling method; they were used only in "bridge" assessments, discussed in Section 3.4.) A bundle size of 29 accommodated the number of students in a typical assessment session plus an allowance for additional students. The bundle size also ensured that each booklet would appear equally often in each position of a bundle, which improved the chance that each booklet would be used the same number of times throughout the assessment.

For grade 3/age 9, 46 booklets numbered 6 through 51 were bundled as follows:

Bundle #1 contained booklets 6 through 34
Bundle #2 contained booklets 35 through 51, then 6 through 17
Bundle #3 contained booklets 18 through 46
Bundle #4 contained booklets 47 through 51, then 6 through 29

This process continued through 46 bundles, at which point booklet 6 was again the first booklet in the bundle, and another 46-bundle cycle was started. Within a cycle, every booklet appeared exactly once in each of the 29 bundle positions.

For grade 7/age 13, 62 booklets numbered 6 through 67 were bundled in the same manner. The cycle contained 62 different bundles. Again, within a cycle, every booklet appeared exactly once in each of the 29 bundle positions.

The bundling procedure differed slightly for grade 11/age 17 in order to achieve a larger sample size for the assessment of literature and U. S. history. The four literature and U.S. history booklets, numbered 92 through 95, were included seven times more often than the other booklets in the bundling cycle. Booklets 92 through 95 were placed as evenly as possible among the other 86 booklets, numbered 6 through 91. Note the appearance of booklets 92 through 95 at the beginning of the first bundle:

6 7 8 92 9 10 11 93 12 13 14 94 15 16 17 95 18 19 20 92 21 22 23 93...

The grade 11/age 17 cycle contained 114 different bundles. Within a cycle, booklets 6 through 91 appeared exactly once in each of the 29 bundle positions; booklets 92 through 95 appeared exactly seven times.

The bundling cycle was repeated until enough booklets had been bundled to satisfy the spiral sample sizes targeted for each grade/age, listed in Table 3-8 at the end of the chapter.

3.4 BRIDGE ASSESSMENT INSTRUMENTS (Booklets 1-5)

At each grade/age, booklets 1 to 5 are "bridge" instruments, used to measure the effects of changes in

- Student age definition (from calendar year to school year)
- Time of testing (from a schedule that is staggered across fall, winter, and spring to a schedule that tests in spring only)
- Mode of administration (from administrations in which students have items read to them from a tape recorder to administrations in which students read items on their own)

These bridge instruments are so called because they build a bridge between the old and new methods. Two separate bridge assessments, Bridge A and Bridge B, were conducted to measure the effects of the three changes. These assessments were conducted at the age level (age 9, age 13, and age 17) only; age 17 was not included in one of the bridge assessments.

3.4.1 Change in Age Definition and Time of Testing (Bridge A: Booklets 1-3)

In 1985-86, NAEP changed the age definition and the time of testing for 9-year-olds and 13-year-olds. The effects of these changes were measured by Bridge A.

In previous assessments, 9- and 13-year-olds were defined as those born during the appropriate calendar year preceding the assessment; 17-year-olds were defined as those born between October 1 and September 30 of the appropriate years preceding the assessment. In the 1985-86 assessment, all three age levels were defined by the October 1 to September 30 interval, both to attain comparability and to link the birth cohorts more closely to school entrance age requirements. With the revised age definitions, the modal grades (the grades into which most students of particular ages fall) are now 3, 7, and 11, which are uniformly four years apart.

Each age level had formerly been assessed at a different time during the school year: 13-year-olds in October and November, 9-year-olds in January and February, and 17-year-olds in March and April. In 1985-86, NAEP began assessing all students in the spring, to eliminate variation in the time of testing and to coordinate the assessment with near-completion of the curriculum year.

Booklets 1, 2, and 3 at ages 9 and 13 contain mathematics, science, reading, and common background blocks. The weighted samples of students who received each booklet, when combined with the weighted samples of students excluded from participation, are each equivalent to the sample of the national student population (see Chapter 4 for more information).

Because computer competence was assessed for the first time in 1985-86, the requirement for bridging in this learning area did not apply. Since neither age definition nor time of testing was changed for 17-year-olds, those students were not included in the Bridge A assessments.

Bridge A assessments were conducted for 9-year-olds from January 6 to January 31, 1986 and for 13-year-olds from November 4 to December 13, 1985.

3.4.2 Change in Mode of Administration (Bridge B: Booklets 4-5)

Prior to the 1983-84 assessment, items were presented aurally and their timing was paced through the use of a tape recorder. The method of spiralling booklets, introduced in 1983-84, requires administration of items by printed page alone.

To permit links between the data from both old and new administration procedures, the Bridge B assessment was conducted for ages 9, 13, and 17 in the learning areas of mathematics and science. Reading was not included in this study because a bridge assessment of reading was conducted in 1983-84; computer competence and U.S. history were not included because they were not assessed under the old procedures; and literature was not included because no links are possible to the data from previous literature assessments.

Booklets 4 and 5 at each age were used for Bridge B. These booklets contain mathematics, science and common background blocks. The weighted samples of students who received each booklet, when combined with the weighted samples of students excluded from participation, are equivalent to the sample of the national student population (see Chapter 4 for more information).

Bridge B was conducted in separate sessions during the same time period as the spiral sessions, which were administered from February 17 to May 2, 1986.

3.5 QUESTIONNAIRES

In addition to the student assessment booklets, four other instruments were administered to collect data about school characteristics, teachers, excluded students and computer coordinators. The methods through which these populations were sampled are described in Chapter 4. These four instruments, which were designed in the form of questionnaires, are described in the following paragraphs.

3.5.1 The Teacher Questionnaire*

NAEP gathered information on the curricula and teaching methods used by selected teachers in schools in which spiral assessments were administered. The data were provided by teachers who completed a questionnaire which included questions about years of teaching experience, course curricula, use of classroom time, homework assignments, and teaching materials used.

Included in the sample for grade 3/age 9 were teachers of English or language arts; for grade 7/age 13, teachers of English, mathematics, or science; and for grade 11/age 17, teachers of English, mathematics, science, or U.S. history.

3.5.2 The School Characteristics and Policy Questionnaire

The School Characteristics and Policy Questionnaire was completed by the school principal or his or her representative for each school included in the spiral assessment and all bridge assessments. The questionnaire was used to gather information about school administration, staffing patterns, special programs, subject requirements, and use and availability of computers. Data collected from the School Characteristics and Policy Questionnaire can be found in the school files on each data tape.

---

*Data not available until Version 2.0

3.5.3 The Excluded Student Questionnaire*

This questionnaire was completed by school personnel for every student who was selected for inclusion in the NAEP sample, but who was unable to respond to items because he or she had limited English language proficiency, was educable mentally retarded, functionally disabled or had other difficulties (for example, emotional distress because of a family situation). The questionnaire was used to gather information about special education, language, and other student programs.

3.5.4 The Computer Coordinator Questionnaire

This questionnaire was completed by the computer coordinator, if there was one, for each school included in the spiral assessment and the Bridge B assessments. The questionnaire was used to gather information about subjects aided by computer instruction, computer topics and courses taught, and computer resources available. Data from the Computer Coordinator Questionnaire can be found on the school files of each data tape.

3.6 TABULAR SUMMARY OF ASSESSMENT INSTRUMENTS

The tables on the following pages summarize the major characteristics of the 1985-86 student assessment instruments.

Tables 3-2, 3-3, and 3-4 provide information about item blocks. Included are the numbers of achievement and attitude items in each block, the booklets in which each block appears, and the number of times each block was used in booklets.

Tables 3-5, 3-6, and 3-7 show which blocks were contained in each booklet. Table 3-8 presents the target sample sizes used to determine the required numbers of assessment blocks and booklets.

Tables 3-9, 3-10, and 3-11 are matrices showing the number of times each block is paired with every other block in the spiral sample.

Tables 3-12, 3-13, and 3-14 show the actual numbers of students who were administered each spiral block and bridge booklet. These tables also include corresponding weighted numbers, standard errors, and coefficients of variation.

---

*Data not available until Version 2.0

Table 3-2

Block Information, Grade 3/Age 9

BLOCK TYPE	BLOCK ID	NUMBER OF ACHIEVEMENT ITEMS	NUMBER OF ATTITUDE ITEMS	APPEARS IN BOOKLETS	NUMBER OF OCCURRENCES
Common Background	BA9	-	28	1 through 51	51
Reading	9R$_{1}$	10	11	1 7 12 17 26 37 48	7
	9R$_{2}$	11	9	2 10 35 38 43 45 48	7
	9R$_{3}$	10	7	3 10 21 23 26 39 51	7
	9R$_{4}$	12	4	13 24 25 48 49 51	6
	9R$_{5}$	13	11	7 32 35 44 47 51	6
	9R$_{6}$	13	4	14 15 16 26 35 49	6
Mathematics	9M$_{1}$	26	0	1 4 11 15 19 24 36 47	8
	9M$_{2}$	26	0	3 4 11 16 22 28 30 32	8
	9M$_{3}$	16	3	2 5 13 19 20 22 23 46	8
	9M$_{4}$	21	7	11 12 29 39 46 50	6
	9M$_{5}$	17	11	17 21 22 29 31 36	6
	9M$_{6}$	20	8	9 19 25 28 29 38	6
	9M$_{7}$	18	10	14 28 34 36 45 46	6
Science	9S$_{1}$	18	5	1 5 6 16 20 27 41 47	8
	9S$_{2}$	25	0	2 5 14 18 39 40 41 44	8
	9S$_{3}$	20	11	3 4 6 8 9 18 24 37	8
	9S$_{4}$	14	9	6 12 21 30 40 42	6
	9S$_{5}$	15	4	8 23 25 27 34 40	6
	9S$_{6}$	15	4	17 18 27 38 42 50	6
	9S$_{7}$	14	7	8 31 32 41 42 43	6
Computer Competence	9C$_{1}$	20	20	10 13 33 34 37 50	6
	9C$_{2}$	19	15	7 9 15 20 33 43	6
	9C$_{3}$	20	13	30 31 33 44 45 49	6

Table 3-3
Block Information, Grade 7/Age 13

BLOCK TYPE	BLOCK ID	NUMBER OF ACHIEVEMENT ITEMS	NUMBER OF ATTITUDE ITEMS	APPEARS IN BOOKLETS	NUMBER OF OCCURRENCES
Common Background	BA13	-	30	1 through 67	67
Reading	13R$_1$	12	19	1 6 8 24 26 32 67	7
	13R$_2$	10	9	2 18 32 34 36 42 51	7
	13R$_3$	13	15	3 6 16 17 20 36 39	7
	13R$_4$	14	7	12 19 20 24 34 35	6
	13R$_5$	12	6	6 14 30 34 47 60	6
	13R$_6$	13	5	14 20 32 46	6
Mathematics	13M$_1$	37	14	1 4 13 26 29 44 57 59	8
	13M$_2$	37	7	3 4 10 13 37 42 43 50 56	9
	13M$_3$	24	8	2 5 7 13 25 33 35 63	8
	13M$_4$	29	14	10 28 33 39 40 44	6
	13M$_5$	26	17	8 10 25 38 55 59	6
	13M$_6$	36	12	29 33 38 46 47 56	6
	13M$_7$	39	16	7 17 28 53 56 59 65	7
	13M$_8$	43	15	22 25 28 29 30 43	6
	13M$_9$	41	16	7 38 43 44 48 51 65	7
Science	13S$_1$	25	11	1 5 40 49 57 61 62 66	8
	13S$_2$	31	9	2 5 11 30 52 54 62 63	8
	13S$_3$	27	9	3 4 15 21 37 52 53 61 67	9
	13S$_4$	18	9	11 21 31 42 55 66	6
	13S$_5$	18	1	15 31 39 46 58 62	6
	13S$_6$	18	10	16 31 41 48 49 52	6
	13S$_7$	18	9	15 35 41 54 60 65 66	7
	13S$_8$	18	14	8 11 12 37 41 58 61	7
	13S$_9$	18	13	21 22 47 49 54 58	6
Computer Competence	13C$_1$	21	21	9 24 45 51 60 64	6
	13C$_2$	20	15	12 16 23 27 45 53	6
	13C$_3$	24	4	18 22 26 27 63 64	6
	13C$_4$	30	7	9 19 27 46 48 50	6
	13C$_5$	20	10	9 14 17 18 23 40	6
	13C$_6$	21	12	19 23 36 55 64 67	6

Table 3-4
Block Information, Grade 11/Age 17

BLOCK TYPE	BLOCK ID	NUMBER OF ACHIEVEMENT ITEMS	NUMBER OF ATTITUDE ITEMS	APPEARS IN BOOKLETS	NUMBER OF OCCURRENCES
Common Background	BA17	-	49	4 through 95	92
Reading	13R$_1$	12	19	34 38 47 50 54 61 89	7
	13R$_2$	10	9	14 21 26 28 36 81 89	7
	13R$_3$	13	15	13 42 50 81 85 86 90	7
	13R$_4$	14	7	7 9 33 61 79 81 88 92 93 94 95	7
	13R$_5$	12	6	24 26 61 62 75 77 85	7
	13R$_6$	13	5	7 22 35 51 68 85 89	7
Mathematics	17M$_1$	35	14	4 10 31 32 36 40 44 57 64	9
	17M$_2$	35	14	4 12 14 15 20 30 31 59 76	9
	17M$_3$	24	11	5 31 41 43 60 67 68 73 83	9
	13M$_4$	29	14	14 15 25 37 53 64 73 87	8
	13M$_5$	26	17	9 11 40 43 45 53 74 76	8
	17M$_6$	36	10	16 20 24 25 40 41 72	7
	17M$_7$	37	16	9 29 30 32 42 58 72 73	8
	17M$_8$	37	15	6 10 11 34 59 60 72 87	8
	17M$_9$	41	20	25 29 36 37 43 46 59 65	8
	17M$_{10}$	36	10	11 13 20 29 57 64 67 83	8
	17M$_{11}$	37	11	6 13 32 41 46 76 78 87	8
Science	17S$_1$	27	11	5 19 47 71 78 79 82 84 91	9
	17S$_2$	32	9	5 8 18 49 57 62 80 91	8
	17S$_3$	23	9	4 8 44 45 48 63 68 70 82	9
	17S$_4$	20	11	6 16 28 55 56 62 63 71	8
	13S$_5$	18	1	8 17 34 39 56 74 84 86	8
	13S$_6$	18	10	12 16 23 39 65 80 82 88	8
	17S$_7$	20	17	15 44 48 51 52 55 80 84	8
	17S$_8$	20	13	17 23 51 58 63 67 69 91	8
	17S$_9$	20	17	12 18 23 52 58 70 71 86	8
	17S$_{10}$	20	15	18 19 48 56 69 74 78 88	8
	17S$_{11}$	20	9	37 39 47 49 55 60 69 70	8

Table 3-4
(continued)

BLOCK TYPE	BLOCK ID	NUMBER OF ACHIEVEMENT ITEMS	NUMBER OF ATTITUDE ITEMS	APPEARS IN BOOKLETS	NUMBER OF OCCURRENCES
Computer Competence	$17C_1$	23	21	7 21 30 38 65 66 77	7
	$17C_2$	21	15	19 27 33 35 42 53 77	7
	$17C_3$	24	4	22 24 27 38 49 52 90	7
	$17C_4$	17	23	10 26 35 45 54 66 90	7
	$17C_5$	24	20	22 28 33 46 50 66 75	7
	$17C_6$	17	19	17 21 27 54 75 79 83	7
U.S. History	$17H_1$	36	25	92	1
	$17H_2$	36	25	93	1
	$17H_3$	35	25	94	1
	$17H_4$	34	25	95	1
Literature	$17L_1$	30	42	92	1
	$17L_2$	31	42	93	1
	$17L_3$	30	42	94	1
	$17L_4$	30	42	95	1

Table 3-5

Booklet Content, Grade 3/Age 9

BOOKLET	BLOCKS			BOOKLET	BLOCKS		
*1)	$9R_1$	$9M_1$	$9S_1$	27)	$9S_1$	$9S_5$	$9S_6$
*2)	$9S_2$	$9R_2$	$9M_3$	28)	$9M_6$	$9M_2$	$9M_7$
*3)	$9M_2$	$9S_3$	$9R_3$	29)	$9M_4$	$9M_6$	$9M_5$
**4)	$9M_1$	$9M_2$	$9S_3$	30)	$9S_4$	$9C_3$	$9M_2$
**5)	$9S_1$	$9S_2$	$9M_3$	31)	$9S_7$	$9C_3$	$9M_5$
6)	$9S_3$	$9S_4$	$9S_1$	32)	$9R_5$	$9M_2$	$9S_7$
7)	$9R_5$	$9R_1$	$9C_2$	33)	$9C_2$	$9C_1$	$9C_3$
8)	$9S_3$	$9S_7$	$9S_5$	34)	$9S_5$	$9M_7$	$9C_1$
9)	$9C_2$	$9M_6$	$9S_3$	35)	$9R_2$	$9R_5$	$9R_6$
10)	$9R_3$	$9R_2$	$9C_1$	36)	$9M_5$	$9M_7$	$9M_1$
11)	$9M_4$	$9M_1$	$9M_2$	37)	$9C_1$	$9R_1$	$9S_3$
12)	$9R_1$	$9M_4$	$9S_4$	38)	$9M_6$	$9S_6$	$9R_2$
13)	$9R_4$	$9C_1$	$9M_3$	39)	$9S_2$	$9R_3$	$9M_4$
14)	$9S_2$	$9R_6$	$9M_7$	40)	$9S_5$	$9S_2$	$9S_4$
15)	$9M_1$	$9C_2$	$9R_6$	41)	$9S_7$	$9S_1$	$9S_2$
16)	$9M_2$	$9R_6$	$9S_1$	42)	$9S_4$	$9S_6$	$9S_7$
17)	$9S_6$	$9M_5$	$9R_1$	43)	$9R_2$	$9S_7$	$9C_2$
18)	$9S_6$	$9S_3$	$9S_2$	44)	$9C_3$	$9S_2$	$9R_5$
19)	$9M_1$	$9M_6$	$9M_3$	45)	$9M_7$	$9C_3$	$9R_2$
20)	$9S_1$	$9C_2$	$9M_3$	46)	$9M_7$	$9M_4$	$9M_3$
21)	$9M_5$	$9S_4$	$9R_3$	47)	$9M_1$	$9S_1$	$9R_5$
22)	$9M_2$	$9M_5$	$9M_3$	48)	$9R_1$	$9R_2$	$9R_4$
23)	$9R_3$	$9S_5$	$9M_3$	49)	$9R_6$	$9R_4$	$9C_3$
24)	$9R_4$	$9S_3$	$9M_1$	50)	$9C_1$	$9M_4$	$9S_6$
25)	$9R_4$	$9S_5$	$9M_6$	51)	$9R_5$	$9R_4$	$9R_3$
26)	$9R_6$	$9R_3$	$9R_1$				

* Booklet used for Bridge A assessment only
** Booklet used for Bridge B assessment only

Table 3-6

Booklet Content, Grade 7/Age 13

BOOKLET	BLOCKS			BOOKLET	BLOCKS		
*1)	$13R_1$	$13M_1$	$13S_1$	35)	$13R_4$	$13S_7$	$13M_3$
*2)	$13S_2$	$13R_2$	$13M_3$	36)	$13C_6$	$13R_3$	$13R_2$
*3)	$13M_2$	$13S_3$	$13R_3$	37)	$13S_3$	$13M_2$	$13S_8$
**4)	$13M_1$	$13M_2$	$13S_3$	38)	$13M_5$	$13M_6$	$13M_9$
**5)	$13S_1$	$13S_2$	$13M_3$	39)	$13S_5$	$13M_4$	$13R_3$
6)	$13R_3$	$13R_1$	$13R_5$	40)	$13C_5$	$13S_1$	$13M_4$
7)	$13M_7$	$13M_9$	$13M_3$	41)	$13S_7$	$13S_6$	$13S_8$
8)	$13S_8$	$13M_5$	$13R_1$	42)	$13R_2$	$13M_2$	$13S_4$
9)	$13C_1$	$13C_5$	$13C_4$	43)	$13M_2$	$13M_8$	$13M_9$
10)	$13M_4$	$13M_2$	$13M_5$	44)	$13M_4$	$13M_1$	$13M_9$
11)	$13S_8$	$13S_2$	$13S_4$	45)	$13C_2$	$13C_1$	$13R_6$
12)	$13S_8$	$13C_2$	$13R_4$	46)	$13S_5$	$13M_6$	$13C_4$
13)	$13M_2$	$13M_1$	$13M_3$	47)	$13R_5$	$13S_9$	$13M_6$
14)	$13R_6$	$13R_5$	$13C_5$	48)	$13C_4$	$13S_6$	$13M_9$
15)	$13S_3$	$13S_5$	$13S_7$	49)	$13S_6$	$13S_9$	$13S_1$
16)	$13R_3$	$13C_2$	$13S_6$	50)	$13C_4$	$13R_6$	$13M_2$
17)	$13M_7$	$13R_3$	$13C_5$	51)	$13C_1$	$13R_2$	$13M_9$
18)	$13C_5$	$13C_3$	$13R_2$	52)	$13S_6$	$13S_3$	$13S_2$
19)	$13R_4$	$13C_4$	$13C_6$	53)	$13C_2$	$13S_3$	$13M_7$
20)	$13R_6$	$13R_4$	$13R_3$	54)	$13S_7$	$13S_2$	$13S_9$
21)	$13S_9$	$13S_4$	$13S_3$	55)	$13S_4$	$13C_6$	$13M_5$
22)	$13C_3$	$13M_8$	$13S_9$	56)	$13M_6$	$13M_7$	$13M_2$
23)	$13C_6$	$13C_5$	$13C_2$	57)	$13M_1$	$13R_6$	$13S_1$
24)	$13R_4$	$13C_1$	$13R_1$	58)	$13S_9$	$13S_8$	$13S_5$
25)	$13M_5$	$13M_8$	$13M_3$	59)	$13M_1$	$13M_5$	$13M_7$
26)	$13M_1$	$13R_1$	$13C_3$	60)	$13R_5$	$13S_7$	$13C_1$
27)	$13C_3$	$13C_4$	$13C_2$	61)	$13S_1$	$13S_8$	$13S_3$
28)	$13M_7$	$13M_4$	$13M_8$	62)	$13S_2$	$13S_1$	$13S_5$
29)	$13M_6$	$13M_8$	$13M_1$	63)	$13S_2$	$13C_3$	$13M_3$
30)	$13M_8$	$13R_5$	$13S_2$	64)	$13C_1$	$13C_6$	$13C_3$
31)	$13S_4$	$13S_5$	$13S_6$	65)	$13S_7$	$13M_7$	$13M_9$
32)	$13R_1$	$13R_2$	$13R_6$	66)	$13S_1$	$13S_4$	$13S_7$
33)	$13M_4$	$13M_6$	$13M_3$	67)	$13R_1$	$13S_3$	$13C_6$
34)	$13R_2$	$13R_4$	$13R_5$				

* Booklet used for Bridge A assessment only
** Booklet used for Bridge B assessment only

Table 3-7
Booklet Content, Grade 11/Age 17

BOOKLET	BLOCKS			BOOKLET	BLOCKS		
1)	Not Used			49)	$17C_3$	$17S_2$	$17S_{11}$
2)	Not Used			50)	$13R_1$	$13R_3$	$17C_5$
3)	Not Used			51)	$17S_7$	$17S_8$	$13R_6$
*4)	$17M_1$	$17M_2$	$17S_3$	52)	$17C_3$	$17S_9$	$17S_7$
*5)	$17S_1$	$17S_2$	$17M_3$	53)	$13M_5$	$13M_4$	$17C_2$
6)	$17S_4$	$17M_{11}$	$17M_8$	54)	$17C_4$	$17C_6$	$13R_1$
7)	$13R_6$	$13R_4$	$17C_1$	55)	$17S_7$	$17S_{11}$	$17S_4$
8)	$17S_3$	$13S_5$	$17S_2$	56)	$13S_5$	$17S_{10}$	$17S_4$
9)	$13R_4$	$13M_5$	$17M_7$	57)	$17M_1$	$17S_2$	$17M_{10}$
10)	$17M_8$	$17M_1$	$17C_4$	58)	$17M_7$	$17S_8$	$17S_9$
11)	$13M_5$	$17M_{10}$	$17M_8$	59)	$17M_2$	$17M_8$	$17M_9$
12)	$17S_9$	$17M_2$	$13S_6$	60)	$17M_8$	$17S_{11}$	$17M_3$
13)	$17M_{11}$	$13R_3$	$17M_{10}$	61)	$13R_1$	$13R_5$	$13R_4$
14)	$13R_2$	$13M_4$	$17M_2$	62)	$17S_2$	$17S_4$	$13R_5$
15)	$17M_2$	$17S_7$	$13M_4$	63)	$17S_4$	$17S_8$	$17S_3$
16)	$13S_6$	$17S_4$	$17M_6$	64)	$17M_{10}$	$13M_4$	$17M_1$
17)	$17S_8$	$13S_5$	$17C_6$	65)	$13S_6$	$17C_1$	$17M_9$
18)	$17S_{10}$	$17S_9$	$17S_2$	66)	$17C_4$	$17C_1$	$17C_5$
19)	$17S_1$	$17S_{10}$	$17C_2$	67)	$17S_8$	$17M_{10}$	$17M_3$
20)	$17M_2$	$17M_6$	$17M_{10}$	68)	$17S_3$	$13R_6$	$17M_3$
21)	$17C_6$	$13R_2$	$17C_1$	69)	$17S_{11}$	$17S_{10}$	$17S_8$
22)	$17C_5$	$17C_3$	$13R_6$	70)	$17S_3$	$17S_{11}$	$17S_9$
23)	$17S_9$	$13S_6$	$17S_8$	71)	$17S_9$	$17S_4$	$17S_1$
24)	$17M_6$	$13R_5$	$17C_3$	72)	$17M_8$	$17M_7$	$17M_6$
25)	$17M_6$	$13M_4$	$17M_9$	73)	$13M_4$	$17M_7$	$17M_3$
26)	$13R_5$	$13R_2$	$17C_4$	74)	$13S_5$	$13M_5$	$17S_{10}$
27)	$17C_2$	$17C_6$	$17C_3$	75)	$17C_6$	$17C_5$	$13R_5$
28)	$13R_2$	$17C_5$	$17S_4$	76)	$17M_{11}$	$17M_2$	$13M_5$
29)	$17M_7$	$17M_{10}$	$17M_9$	77)	$17C_1$	$17C_2$	$13R_5$
30)	$17M_7$	$17M_2$	$17C_1$	78)	$17M_{11}$	$17S_1$	$17S_{10}$
31)	$17M_1$	$17M_2$	$17M_3$	79)	$13R_4$	$17C_6$	$17S_1$
32)	$17M_1$	$17M_7$	$17M_{11}$	80)	$17S_2$	$13S_6$	$17S_7$
33)	$17C_5$	$17C_2$	$13R_4$	81)	$13R_3$	$13R_4$	$13R_2$
34)	$17M_8$	$13R_1$	$13S_5$	82)	$13S_6$	$17S_3$	$17S_1$
35)	$17C_2$	$17C_4$	$13R_6$	83)	$17M_{10}$	$17C_6$	$17M_3$
36)	$17M_1$	$13R_2$	$17M_9$	84)	$17S_1$	$17S_7$	$13S_5$
37)	$17S_{11}$	$13M_4$	$17M_9$	85)	$13R_5$	$13R_6$	$13R_3$
38)	$17C_1$	$17C_3$	$13R_1$	86)	$13S_5$	$17S_9$	$13R_3$
39)	$17S_{11}$	$13S_5$	$13S_6$	87)	$13M_4$	$17M_8$	$17M_{11}$
40)	$13M_5$	$17M_6$	$17M_1$	88)	$17S_{10}$	$13S_6$	$13R_4$
41)	$17M_6$	$17M_{11}$	$17M_3$	89)	$13R_6$	$13R_1$	$13R_2$
42)	$13R_3$	$17C_2$	$17M_7$	90)	$17C_4$	$17C_3$	$13R_3$
43)	$13M_5$	$17M_9$	$17M_3$	91)	$17S_2$	$17S_1$	$17S_8$
44)	$17M_1$	$17S_7$	$17S_3$	92)	$13R_4$	$17H_1$	$17L_1$
45)	$17S_3$	$17C_4$	$13M_5$	93)	$17H_2$	$13R_4$	$17L_2$
46)	$17C_5$	$17M_{11}$	$17M_9$	94)	$17L_3$	$13R_4$	$17H_3$
47)	$17S_1$	$17S_{11}$	$13R_1$	95)	$17L_4$	$17H_4$	$13R_4$
48)	$17S_7$	$17S_3$	$17S_{10}$				

*Booklet used for Bridge B assessment only

Table 3-8

Target Sample Sizes by Grade/Age[1]

	Grade 3/Age 9	Grade 7/Age 13	Grade 11/Age 17		Total
Spiral	19,933	26,867	37,267	10,800[2]	94,867
Bridge	10,000	10,000	4,000	-- [2]	24,000
Total	29,933	36,867	41,267	10,800[2]	118,867

[1] Actual sample sizes are given in Table 4-2 in Chapter 4

[2] Literature and U.S. history sample

Table 3-9

Block-to-Block Occurrence Matrix, Grade 3/Age 9 Spiral Booklets

	R1	R2	R3	R4	R5	R6	M1	M2	M3	M4	M5	M6	M7	S1	S2	S3	S4	S5	S6	S7	C1	C2	C3
R1	6	1	1	1	1	1				1	1					1	1		1		1	1	
R2		6	1	1	1	1						1	1						1	1	1	1	1
R3			6	1	1	1			1	1	1				1		1	1			1		
R4				6	1	1	1		1			1				1		1			1		1
R5					6	1	1	1						1	1					1		1	1
R6						6	1	1					1	1	1							1	1
M1							6	1	1	1	1	1	1	1		1						1	
M2								6	1	1	1	1	1	1			1			1			1
M3									6	1	1	1	1	1				1			1	1	
M4										6	1	1	1		1		1		1		1		
M5											6	1	1				1		1	1			1
M6												6	1			1		1	1			1	
M7													6		1			1			1		1
S1														6	1	1	1	1	1	1		1	
S2															6	1	1	1	1	1			1
S3																6	1	1	1	1	1	1	
S4																	6	1	1	1			1
S5																		6	1	1	1		
S6																			6	1	1		
S7																				6		1	1
C1																					6	1	1
C2																						6	1
C3																							6

Table 3-10

Block-to-Block Occurrence Matrix, Grade 7/Age 13 Spiral Booklets

	R_1	R_2	R_3	R_4	R_5	R_6	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	C_1	C_2	C_3	C_4	C_5	C_6
R_1	6	1	1	1	1	1	1				1							1					1		1		1			1
R_2		6	1	1	1	1		1							1				1						1		1		1	1
R_3			6	1	1	1				1			1							1	1					1			1	1
R_4				6	1	1			1													1	1		1	1		1		1
R_5					6	1						1		1			1					1		1	1				1	
R_6						6	1	1								1									1	1		1	1	
M_1							6	1	1	1	1	1	1	1	1	1											1			
M_2								7	1	1	1	1	1	1	1			1	1				1					1		
M_3									6	1	1	1	1	1	1		1					1					1			
M_4										6	1	1	1	1	1	1				1									1	
M_5											6	1	1	1	1				1				1							1
M_6												6	1	1	1					1				1				1		
M_7													7	1	2			1				1				1			1	
M_8														6	1		1							1			1			
M_9															7						1	1			1			1		
S_1																6	1	1	1	1	1	1	1	1					1	
S_2																	6	1	1	1	1	1	1	1			1			
S_3																		7	1	1	1	1	2	1		1				1
S_4																			6	1	1	1	1	1						1
S_5																				6	1	1	1	1				1		
S_6																					6	1	1	1		1		1		
S_7																						7	1	1	1					
S_8																							7	1		1				
S_9																								6			1			
C_1																									6	1	1	1	1	1
C_2																										6	1	1	1	1
C_3																											6	1	1	1
C_4																												6	1	1
C_5																													6	1
C_6																														6

Table 3-11

Block-to-Block Occurrence Matrix, Grade 11/Age 17 Spiral Booklets*

	R1	R2	R3	R4	R5	R6	M1	M2	M3	M4	M5	M6	M7	M8	M9	M;	M:	S1	S2	S3	S4	S5	S6	S7	S8	S9	S;	S:	C1	C2	C3	C4	C5	C6
R1	7	1	1	1	1	1								1				1				1						1	1		1	1	1	1
R2		7	1	1	1	1	1	1		1					1						1								1			1	1	1
R3			7	1	1	1							1			1	1					1				1				1	1	1	1	
R4				7	1	1					1		1					1					1				1		1	1			1	1
R5					7	1						1							1		1								1	1	1	1	1	1
R6						7			1											1				1	1				1	1	1	1	1	
M1							8	1	1	1	1	1	1	1	1	2	1		1	1				1								1		
M2								8	1	2	1	1	1	1	1	1	1						1	1		1			1					
M3									8	1	1	1	1	1	1	2	1			1					1			1						1
M4										8	1	1	1	1	2	1	1							1				1		1				
M5											8	1	1	1	1	1	1			1		1					1			1		1		
M6												7	1	1	1	1	1				1		1								1			
M7													8	1	1	1	1								1	1			1	1				
M8														8	1	1	2				1	1						1				1		
M9															8	1	1						1					1	1				1	
M;																8	1		1						1									1
M:																	8	1			1						1						1	
S1																		8	1	1	1	1	1	1	1	1	2	1		1				1
S2																			7	1	1	1	1	1	1	1	1	1			1			
S3																				8	1	1	1	2	1	1	1	1				1		
S4																					8	1	1	1	1	1	1	1					1	
S5																						8	1	1	1	1	2	1						1
S6																							8	1	1	2	1	1	1					
S7																								8	1	1	1	1			1			
S8																									8	2	1	1						1
S9																										8	1	1			1			
S;																											8	1		1				
S:																												8			1			
C1																													7	1	1	1	1	1
C2																														7	1	1	1	1
C3																															7	1	1	1
C4																																7	1	1
C5																																	7	1
C6																																		7

*Semicolon (;) represents 10; colon (:) represents 11.

Table 3-12

Number of Students Administered Each Spiral Block and Bridge Booklet Grade 3/Age 9

Spiral Blocks (Grade 3/Age 9)(Booklets 6-51)

Block	Total	Weighted N	Standard Error (of Weighted N)	Coefficient of Variation
9R1	2778	519199	5066	1.0
9R2	2778	514795	7812	1.5
9R3	2748	504935	5666	1.1
9R4	2771	513876	5933	1.2
9R5	2779	518050	6879	1.3
9R6	2788	520851	5532	1.1
9M1	2782	520673	6204	1.2
9M2	2796	515681	7841	1.5
9M3	2754	513679	9010	1.8
9M4	2762	501816	6684	1.3
9M5	2742	505693	9042	1.8
9M6	2764	507773	6551	1.3
9M7	2796	518787	6318	1.2
9S1	2800	519771	6540	1.3
9S2	2796	513036	7704	1.5
9S3	2774	517181	6453	1.2
9S4	2774	512470	6152	1.2
9S5	2783	506764	7474	1.5
9S6	2789	510028	6170	1.2
9S7	2810	515453	7194	1.4
9C1	2770	502409	7700	1.5
9C2	2754	509985	6458	1.3
9C3	2773	513073	7548	1.5
Total Spiral	21287	3931992	18966	0.5

Bridge Booklets (Age 9)

Booklet	Total	Weighted N	Standard Error (of Weighted N)	Coefficient of Variation
1	2315	3098639	16593	0.5
2	2361	3104555	20282	0.7
3	2256	3112834	14390	0.5
4	1994	3151352	20051	0.6
5	2048	3121844	26874	0.9

Table 3-13

Number of Students Administered Each Spiral Block and Bridge Booklet Grade 7/Age 13

Spiral Blocks (Grade 7/Age 13)(Booklets 6-67)

Block	Total	Weighted N	Standard Error (of Weighted N)	Coefficient of Variation
13R1	2645	387896	5534	1.4
13R2	2688	395605	7477	1.9
13R3	2680	385111	6283	1.6
13R4	2695	395021	7079	1.8
13R5	2679	389566	5566	1.4
13R6	2692	394547	6344	1.6
13M1	2678	387547	4521	1.2
13M2	3130	452148	6003	1.3
13M3	2661	392482	5393	1.4
13M4	2723	393767	6091	1.5
13M5	2696	381586	5257	1.4
13M6	2695	386437	6302	1.6
13M7	3092	443077	5663	1.3
13M8	2697	390203	6705	1.7
13M9	3125	452549	7000	1.5
13S1	2695	389658	6989	1.8
13S2	2660	387668	6189	1.6
13S3	3125	451104	4213	0.9
13S4	2682	389091	4545	1.2
13S5	2670	383151	3623	0.9
13S6	2670	392571	5429	1.4
13S7	3109	445846	6180	1.4
13S8	3115	447356	6336	1.4
13S9	2676	381161	6605	1.7
13C1	2665	389025	4492	1.2
13C2	2672	387848	5357	1.4
13C3	2667	390099	4732	1.2
13C4	2677	388399	5812	1.5
13C5	2688	392545	8021	2.0
13C6	2657	380658	4698	1.2
Total Spiral	27668	4007907	16317	0.4

Bridge Booklets (Age 13)

Booklet	Total	Weighted N	Standard Error (of Weighted N)	Coefficient of Variation
1	2075	2937402	21332	0.7
2	2054	2950983	24449	0.8
3	2071	2943837	21182	0.7
4	2032	3008026	22738	0.8
5	2146	3028806	15282	0.5

Table 3-14

Number of Students Administered Each Spiral Block and Bridge Booklet
Grade 11/Age 17

Spiral Blocks (Grade 11/Age 17)(Booklets 6-95)

Block	Total	Weighted N	Standard Error (of Weighted N)	Coefficient of Variation
13R1	2449	252747	4142	1.6
13R2	2428	253214	4501	1.8
13R3	2395	250738	3766	1.5
13R4	12192	1264606	6608	0.5
13R5	2436	254622	3928	1.5
13R6	2417	247014	3840	1.6
17M1	2823	295171	4062	1.4
17M2	2795	289562	4933	1.7
17M3	2782	290965	5121	1.8
13M4	2812	291714	2968	1.0
13M5	2788	290699	4275	1.5
17M6	2443	252921	4624	1.8
17M7	2786	294005	3644	1.2
17M8	2786	292137	4173	1.4
17M9	2798	293414	5479	1.9
17M10	2785	291233	3098	1.1
17M11	2780	291096	3872	1.3
17S1	2747	284402	5606	2.0
17S2	2421	250694	3077	1.2
17S3	2828	296071	5369	1.8
17S4	2835	298942	4467	1.5
13S5	2772	286892	4105	1.4
13S6	2771	288266	4483	1.6
17S7	2814	290561	4111	1.4
17S8	2818	293055	3566	1.2
17S9	2785	293245	4886	1.7
17S10	2751	286728	3738	1.3
17S11	2851	299574	6761	2.3
17C1	2433	253647	3409	1.3
17C2	2457	254457	3547	1.4
17C3	2444	252893	3757	1.5
17C4	2438	254840	3514	1.4
17C5	2427	254089	3760	1.5
17C6	2424	252106	2431	1.0
17H1	2441	252282	3675	1.5
17H2	2428	251539	2892	1.1
17H3	2459	254264	3037	1.2
17H4	2446	254201	2653	1.0

Table 3-17
(continued)

Spiral Blocks (Grade 11/Age 17)(Booklets 6-95)				
Block	Total	Weighted N	Standard Error (of Weighted N)	Coefficient of Variation
17L1	2441	252282	3675	1.5
17L2	2428	251539	2892	1.1
17L3	2459	254264	3037	1.2
17L4	2446	254201	2653	1.0
Total Spiral	39753	4136965	13025	0.3

Bridge Booklets (Age 17)				
Booklet	Total	Weighted N	Standard Error (of Weighted N)	Coefficient of Variation
4	1934	3240017	13772	0.4
5	1934	3252949	10159	0.3

Chapter 4

SAMPLE SELECTION

Chapter 4: SAMPLE SELECTION[1]

4.1 INTRODUCTION

This chapter describes the special features of the 1985-86 sample design and the methods used by Westat, Inc., the survey subcontractor, to select the sample. Also discussed in this chapter are the sampling weights provided by Westat and how they were derived. A discussion of how to use these sampling weights can be found in Chapter 7.

4.2 GRADE/AGE DEFINITION

As in the 1983-84 assessment, the 1985-86 assessment sampled students by grade as well as age. This sampling enhances the utility of NAEP data, since school districts traditionally delineate groups of students by grade rather than age. However, the 1985-86 main sample uses a different definition of the ages 9 and 13 than that of prior assessments and assesses these students at a different time of year. In past assessments, 9-year-olds were defined on a calendar year basis and were assessed in the wintertime; 13-year-olds were also defined on a calendar year basis and were assessed in the fall of the school year; 17-year-olds were defined on an October through September basis and were assessed in the spring of the school year.

In 1985-86, uniform age definitions and uniform administration dates were adopted for the main assessment. The students of a particular age were defined as the students born between October 1 and September 30 of the appropriate years preceding the assessment. All ages were assessed in the spring. This modification of the previous age definitions (for ages 9 and 13) had the effect of changing the modal grades to 3, 7, and 11 from 4, 8, and 11, the grades assessed in the past. This change resulted in the need for bridge studies, as detailed below.

The 1985-86 main assessment includes three student cohorts: students who were either in the third grade or 9 years old; students who were either in the seventh grade or 13 years old; and students who were either in the eleventh grade or 17 years old. The main assessment represents two overlapping samples. The first sample represents students of specified grades (who could be of any age). The second sample represents students of specified ages (who could be of any grade). For the age 17 students, this sample is comparable to samples from previous NAEP assessments. Because of the changes in age definitions for age 9 and age 13, the samples for those ages are not directly comparable to previous NAEP assessments.

---

[1] Data collected from the Teacher Questionnaire and the Excluded Student Questionnaire, and sampling weights for schools, excluded students, and teachers associated with students are not contained on the Version 1.0 data tapes, but will be included in Version 2.0.

4.3 BRIDGE SAMPLES

To determine the possible effects of changes in age definitions as well as in mode of administration (elimination of the audiotape used for pacing the exercises), two bridge studies were conducted. The term "bridge" describes the purpose of these studies, which is to provide the statistical linkage between the 1985-86 data and data from previous assessments. The two bridge studies, which were conducted at age levels only, are as follows:

> Bridge A: A bridge study to measure the effect of changing the age definitions and the time of year in which the assessment data were collected. Since these changes affected only ages 9 and 13, a bridge sample was not necessary for age 17. Since trend data have been traditionally collected only by age, grade sampling was unnecessary.
>
> This bridge study addresses the learning areas of reading, mathematics, and science and consists of three booklets each for ages 9 and 13. Within each booklet, mathematics and science blocks were administered using a tape recorder; reading blocks were administered by pencil and paper only. The respondents to any one of the three booklets assigned to a given age comprise a representative sample of the population of all students of that age. Since, for this bridge study, the traditional age definitions were used, each of these samples is comparable to samples from previous NAEP assessments.
>
> Bridge B: A bridge study to measure the effect of changing from tape-recorded to printed administration. The new definitions of age and time of assessment were used. These bridge samples were collected at all three age levels for the learning areas of science and mathematics. At all ages, two bridge booklets were administered. The respondents to any one of the two bridge booklets assigned to a given age constitute a representative sample of the population of all students of that age, where the new age definition is being used. Each of these samples is comparable to the corresponding age sample from the main (spiral) assessment.
>
> The full 1985-86 sample is summarized in Section 4.6.

4.4 SAMPLE DESIGN

The sample of students for the 1985-86 NAEP assessment was selected using a complex multistage sample design involving the sampling of students from selected schools within 94 selected geographic regions, called primary sampling units, from across the United States. The sample design will be described in detail in a technical report to be issued in 1987 by Westat, Inc., the firm subcontracted by ETS to select the sample. This section will provide an overview of the design.

4.4.1 Primary Sampling Units

In the first stage of sampling, the United States was divided into geographic units comprised of counties, or groups of contiguous counties, which met a minimum size requirement. These units are called primary sampling units (PSUs). Twelve subuniverses were then defined as follows:

First, the PSUs were classified by the four regions. In each region, PSUs were classified as MSA (Metropolitan Statistical Area) or non-MSA. In the Southeast and West regions, the PSUs were further classified as high minority (20 percent of the population in the 1980 Census was either Black or Hispanic) or not. The resulting subuniverses are shown in Table 4-1. Among the larger PSUs, 34 were designated as certainty units to be selected with probability one. Within each major stratum (the subuniverses), further stratification was achieved by ordering the noncertainty PSUs according to several additional socioeconomic characteristics, yielding 60 strata--a total of 94 strata in all. One PSU was selected with probability proportional to size from each of the 60 noncertainty strata. PSUs within the high-minority subuniverses were sampled at twice the rate of PSUs in the other subuniverses.

Table 4-1
The Sampling Subuniverses

	NORTHEAST	SOUTHEAST	CENTRAL	WEST
MSA PSUs	1	3 high minority	7	9 high minority
		4 low minority		10 low minority
non-MSA PSUs	2	5 high minority	8	11 high minority
		6 low minority		12 low minority

These PSUs were used for both the spiral assessments and the Bridge B assessments. The Bridge A assessments used a subsample of 64 PSUs which were selected from the complete set of 94 PSUs with probability proportional to a measure of size.

4.4.2 Schools

In the second stage of sampling, all public, private, Catholic, Bureau of Indian Affairs, and Department of Defense schools were listed according to the three grade/age groups within each of the 94 PSUs. High-minority schools were oversampled to ensure adequate sample sizes, thereby enhancing the reliability of estimates.

Schools within each PSU were selected (without replacement) with probabilities proportional to assigned measures of size. Roughly equal measures of size were assigned to schools containing estimates of grade/age-eligible students ranging from 20 to 150 (for grade 3/age 9), or 20 to 200 (for grade 7/age 13 and grade 11/age 17). Schools above the indicated maximum size were selected with probabilities proportional to the number of grade/age-eligible students. Schools with less than 20 estimated grade/age eligibles were assigned considerably lower measures of size, since they had higher per-student administrative costs.

4.4.3 Assigning Sessions to Schools, by Type

The assignment of sessions to schools served as the third stage of sampling. This assignment was done separately by the three types of sessions, designated spiral, Bridge A, and Bridge B, which represent separate samples of the student population.

The Bridge A assessments involved three distinct booklets each for ages 9 and 13 (using the old definitions). Schools to participate in this assessment (conducted in the fall and winter) were selected from the subsample of 64 PSUs which had been designated as the Bridge A PSUs. Each of the three distinct booklets was to be administered once within each of the PSUs. To avoid the possibility that a particular bridge session might be assigned to a school with only one or very few eligibles, small schools were clustered with other schools in the same PSU to form clusters of a specified minimum number of eligibles. Bridge sessions were then assigned within each PSU by selecting a systematic sample of three school clusters at each age with probability proportional to the estimated number of age eligibles within the school (or school cluster).

Schools which were selected for the Bridge A assessment of a grade/age class were excluded from the spiral and Bridge B assessments of the grade/age class.

All remaining selected schools were used for the spiral and Bridge B assessments. For the Bridge B assessments, there were two distinct booklets at each age class, each of which was administered once within each of the 94 PSUs. A systematic sample of two schools within each PSU was selected for the Bridge B assessment. After this selection, spiral sessions were assigned to schools at a rate approximately proportional to the estimated number of eligible students who would be available after accounting for the initial assignment of the Bridge B sessions. With the exception of the smallest schools selected for the Bridge B assessment, each school selected for the combined spiral and Bridge B assessments was allocated at least one spiral assessment.

4.4.4 Sampling Students

In the fourth stage of sampling, a consolidated list of all grade-eligible and age-eligible students was established for each selected school. A systematic selection of eligible students was made and, for the spring assessment, students were assigned by Westat district supervisors to spiral or Bridge B sessions, depending on whether the assessment was to be administered by pencil and paper or tape recorder.

4.4.5 Sampling Excluded Students

Some students selected for the sample were deemed unassessable by the school authorities because they had limited English language proficiency, were judged as being educable mentally retarded, were functionally disabled, or had other difficulties (for example, emotional distress because of a family situation). In these cases, an Excluded Student Questionnaire was filled out by the school staff listing the reason for excluding the student and providing some background information.

4.4.6 Sampling Teachers

The Teacher Questionnaire was administered to the teachers of a subgroup of students sampled for spiral sessions. The purpose of this sample was to estimate the number (proportion) of students whose teachers had various attributes, not to estimate the attributes of the teacher population. Therefore, statements such as, "20 percent of students have teachers who have..." are appropriate in discussing Teacher Questionnaire data, but statements such as, "20 percent of teachers have..." are not. The number of teachers sampled was equal to the number of spiral sessions conducted in the school. The teachers selected taught certain subject areas. One subject was designated for each spiral assessment in the school. Therefore, if a school had two spiral assessments, two subjects were designated, perhaps the same or different subjects. For grade 3/age 9, the subject was always English or language arts. For grade 7/age 13 the subject(s) could have been either English/language arts, mathematics, or science. For grade 11/age 17, the subject(s) could have been either English/language arts, U.S. history, mathematics, or science.

For each spiral session, a subsample of students was selected and the school coordinator was asked to identify, for each selected student, the teacher in the designated subject area who was teaching the student. These instructors completed a Teacher Questionnaire. Please note that since a number of students may have had the same teacher, and some teachers did not complete the questionnaire, the number of students in the subsample for whom teacher information is available is greater than the number of teachers who completed questionnaires in a given school.

4.4.7 Sampling School Characteristics and Computer Coordinators

The School Characteristics and Policy Questionnaire was distributed in every sampled school. It was mailed to the school by Westat prior to the assessment and picked up by the Westat supervisor, then returned to ETS.

In every school, for all three grade/age levels, a special questionnaire was also given to the school's computer coordinator, if the school had one.

The School Characteristics and Policy and Computer Coordinator questionnaires are described in Chapter 3.

A Principal Questionnaire, distributed to each sampled school by Westat prior to the assessment, was used by Westat to determine an estimate of grade/age-eligible students. Some of the data from the Principal Questionnaire has been retained on the data tapes. See Chapter 6 for an explanation of the Principal Questionnaire data variables.

4.5 SAMPLE WEIGHTS

NAEP uses differential sampling rates. Some subpopulations are deliberately oversampled to obtain larger samples of respondents in various reporting groups. As a result of this oversampling, these subpopulations, which may have different characteristics from the rest of the population, are overrepresented in the sample. Analyses which ignore this overrepresentation may be misleading since these subpopulations have unwarranted impact on the results. To account for these differential probabilities of selection, as well as other aspects of the sample and population being surveyed, sampling weights are supplied on the data files so that each respondent will be appropriately represented in data analysis. These weights should be used for all analyses, whether exploratory or confirmatory.

Information about what weights are provided on the public-use data tapes and how they should be used appears in Section 7.2 of Chapter 7. The remainder of the current section will be devoted to a description of the procedures used by NAEP to derive the sampling weights for the assessed students. Readers who do not need detailed technical information about the derivation of weights may disregard Sections 4.5.1 through 4.5.6.

The weight assigned to a particular student reflects two major components of the sample design and the population being surveyed. The first component, the student's base weight, reflects the probability of selection of the student for participation in a particular type of assessment session (i.e., a particular bridge assessment session or a spiral session). The second component reflects adjustments to the base weight to account for nonresponse and to ensure that estimates, based on these sampling weights, of certain subpopulation totals correspond to values reliably known from external (e.g., Census) sources. This latter form of adjustment is known as post-stratification and can reduce the bias resulting from noncoverage as well as nonresponse and may, additionally, reduce sampling variability.

4.5.1 Base Weight

The base weight assigned to a student is the reciprocal of the probability that the student was invited to a particular type of assessment session, that is, a spiral session or a particular bridge assessment session. That probability is the product of four factors:

1) The probability that the PSU was selected;
2) the conditional probability, given the PSU, that the school was selected;
3) the conditional probability, given the sample of schools in a PSU, that the school was allocated the specified type of session; and
4) the conditional probability, given the school, that the student was invited to the specified type of session.

Thus, the base weight for a student may be expressed as the product

$$W_B = PSUWT \cdot SCHWT \cdot SESSWT \cdot STUDWT$$

where PSUWT, SCHWT, SESSWT, and STUDWT are, respectively, the reciprocals of the preceding probabilities.

4.5.2 Adjustment of Base Weights for Nonresponse

The base weight for a student was adjusted by two nonresponse factors; one to adjust for non-cooperating schools and one to adjust for students who were invited to the assessment but did not appear either in the scheduled session or in a makeup session. Thus, the student nonresponse adjusted weight is of the form

$$W_w = W_B \cdot f_1 \cdot f_2 \cdot f_3$$

where W_B is the student base weight, f_1 is a school nonresponse factor, f_2 is a session nonresponse factor, and f_3 is a student nonresponse factor, each computed as described below.

School Nonresponse Adjustment (SCHNRF and SESNR)

School nonresponse factors were computed separately within each PSU for up to two or three classes of schools using as many nonresponse classes as the number of sampled schools in the PSU and nonresponse pattern allowed. However, since it was required that each class contain at least four or five schools, often only one class was identified in the PSU.

For any nonresponse class, c, the school nonresponse factor for spiral sessions is given by

$$f_{1c} = \frac{\sum_{i \epsilon A} W_i G_i}{\sum_{i \epsilon B} W_i G_i}$$

where

W_i = school weight (the reciprocal of the probability of selection of the school conditional on the PSU);

G_i = estimated number of grade-eligible students in school i;

set A consists of the original sample of eligible schools (including refusing schools but not including substitutes); and

set B consists of all cooperating schools (including schools that were substituted for non-cooperating schools).

Note that, for a substitute school, W_i was computed as if the school had been originally selected by the sampling procedure.

The school nonresponse factor, f_1, appears as SCHNRF on the data tapes.

The session nonresponse factor, f_2, appears as SESNRF on the data tapes.

Student Nonresponse Adjustment (STUNRF)

Student nonresponse adjustment factors were computed separately for spiral sessions and for each of the bridge assessment sessions within each PSU.

The student nonresponse factor, f_3, appears as STUNRF on the data tapes.

Nonresponse Adjustment for Students in Bridge Sessions

For each bridge session, t, in a PSU, the nonresponse factor f_{2t} was computed by

$$f_{2t} = \frac{n_t}{n'_t}$$

where

n_t = number of students invited to the particular Bridge session in the PSU; and

n'_t = number of students who completed the session.

Nonresponse Adjustment for Students in Spiral Sessions

For spiral sessions, the student nonresponse adjustment was made separately for two classes of students: those in or above the modal grade for their age and those below the modal age. This is in recognition of the likely differences between students in the two classes both in their assessed abilities and in their likelihood of nonresponse.

The factor for students in class c in a particular PSU was computed by

$$f_{2c} = \frac{\sum_i W_i n_{ic}}{\sum W_i n'_{ic}}$$

where the summations extend over the schools in the PSU and

n_{ic} = number of spiral invited students in school i and student class c;

n'_{ic} = number of spiral tested students in school i and student class c; and

w_i = the reciprocal of the probability of assignment of a student in school i to a spiral session, conditional on the PSU, adjusted for school nonresponse.

4.5.3 Trimming of Weights (WTRIMF)

In a number of cases, students were assigned extremely large weights. One cause of large weights was underestimation of the number of eligible students in some schools leading to inappropriately low probabilities of selection for those schools. Other extremely large weights arose as the result of high levels of nonresponse coupled with low to moderate probabilities of selection.

Students with extremely large weights have an unusually large impact on estimates such as weighted means. Since the variability in weights contributes to the variance of an overall estimate by an approximate factor $1 + V^2$, where V^2 is the relative variance of the weights, a few extremely large weights are likely to produce large sampling variances of the statistics of interest, especially when the large weights are associated with students with atypical performance characteristics.

In such cases, a procedure of trimming the more extreme weights to values somewhat closer to the mean weight was applied. This trimming can increase the accuracy of the resulting survey estimates, substantially reducing V^2 and hence the sampling variance while possibly introducing a small bias.

The weight trimming adjustment factor appears as WTRIMF on the data tapes.

4.5.4 Post-Stratification (PSTRATF)

As in most sample surveys, the respondent weights are random variables which are subject to sampling variability. Even if there were no nonresponse, the respondent weights would at best provide unbiased estimates of the various subgroup proportions. However, since unbiasedness refers to average performance over a conceptually infinite number of replications of the sampling, it is unlikely that any given estimate, based on the achieved sample, will exactly equal the population value. Furthermore, the respondent weights have been adjusted for nonresponse and a number of extreme weights have been reduced in size.

To reduce the mean squared error of estimates using the sampling weights, these weights were further adjusted so that estimated population totals for a number of specified subgroups of the population, based on the sum of weights of students of the specified type, were the same as presumably better estimates derived from other sources. This adjustment, called post-stratification, reduces the mean squared error of estimates relating to student populations that span several subgroups of the population. The post-stratification was done separately for the spiral sessions and each of the bridge sessions within each grade/age group, because each of these can be viewed as separate samples of the appropriate population.

The post-stratification adjustment factor appears as PSTRATF on the data tapes.

4.5.5 The Final Student-Weight: The Full-Sample Weight (WEIGHT)

The final weight assigned to a student is the student full-sample weight. This weight is the student's base weight after the application of the various adjustments described above.

The student full-sample weight appears as WEIGHT on the data tapes and is the product of the variables PSUWT, SCHWT, SCHNRF, SESSWT, SESNR, STUDWT, STUNRF, WTRIMF and PSTRATF. The student full-sample weight should be used to derive all estimates of population and subpopulation characteristics such as the proportion of students of a specified type who would respond in a certain way to an exercise.

4.5.6 Weights Used for Variance Estimation (SRWTxx)

Variances for NAEP estimates are computed by a jackknife repeated replication method. This method involves the use of a series of weights SRWT01 through SRWT38 (for the spiral sample).

For information about the use of these weights and how they were computed, see Section 7.3 in Chapter 7.

4.6 TABULAR SUMMARY OF THE 1985-86 SAMPLES

The 1985-86 NAEP database includes a number of different samples from each grade/age population of students. These samples include:

- The spiral sample of students who are in the grade/age combination and who were, in the school's judgment, assessable
- The two Bridge B samples (bridge booklets 4 and 5) of students who are of the specified age and who were assessable
- The excluded students from the combined spiral assessment and the Bridge B assessments who, in the school's judgment, were not assessable
- The three Bridge A samples (bridge booklets 1, 2 and 3) of students who are of the specified age (using the old age definition) and who were assessable
- The excluded students from the combined Bridge A assessments.

Westat has developed an appropriate set of sampling weights to use for each of these samples. All estimates of population parameters use these weights. Details of the construction of these weights appear in Section 4.5 of this chapter. A discussion of which weights to use for which analyses appears in Chapter 7. This section provides summary statistics for each of the samples including the sample sizes and estimated population totals.

Table 4-2 shows the sizes of the various samples and their sampling weights by grade/age combination and by type of sample. The top position of the table addresses the samples from the spiral and Bridge B assessments.

The sums of the weights for the spiral sample estimate the numbers of grade/age-eligible students in the total population who would be deemed assessable by their schools. Note that this sample includes the respondents to any block used in the spiral assessment of the given grade/age.

The two Bridge B samples, bridge booklet 4 and bridge booklet 5 are defined by age only; each sum of weights estimates the numbers of age-eligible students (using the new age definition) in the total population who would be deemed assessable by their schools. The differences in the estimated total number of age-eligible students from the two Bridge B assessments are due to sampling variance.

Finally, the sums of weights of the excluded students from the combined spiral and Bridge B assessments estimate the numbers of grade/age-eligible students in the total population who, in their school's opinion, would not be assessable.

To estimate the total number of in-school students in a grade/age category, add

- the sum of the weights for the spiral sample

to

- the sum of the weights for the excluded students from the spiral and Bridge B samples

To estimate the total number of students (of any grade) in an age category, add

- the sum of the weights from either of the two Bridge B assessments

to

- the sum of the weights for the age-eligible excluded students from the spiral and Bridge B samples

The lower part of Table 4-2 shows the sample sizes and sum of weights for the three Bridge A samples. These assessments are of the populations of students defined by the old age definitions. Since the age definition for the 17-year-olds was unchanged in the main assessment, no such bridge was necessary. For ages 9 and 13, the sum of weights for each of the Bridge A samples, bridge booklet 1, bridge booklet 2, and bridge booklet 3, provides an estimate of the number of assessable age-eligible students in the total population.

The sum of the weights of the excluded students from the Bridge A assessments estimates the number of age-eligible students in the total population who, in their school's opinion, would have been unassessable.

To estimate the total number of students (of any grade) in an age category using the old age definition, add

- the sum of the weights from any of the three Bridge A assessments

to

- the sum of the weights for the excluded students from the Bridge A assessments

In most cases, the number of students in a grade/age combination is not of interest; a researcher will be interested in estimating the number of students at either a grade or an age. An estimate of the number of students at an age level can be made by summing the weights of only the age-eligible students, and an estimate of the number of students in a grade by summing the weights of grade-eligible students.

Tables 4-3, 4-4, and 4-5 show how many students at each grade level are at, in, or above the modal age for that grade, and how many at each age level are at, in, or above the modal grade for that age. These figures were computed from the spiral sample only. Along with the counts from this sample, the sum of the weights (Weighted N) for each category is presented, and these sums are estimates of the numbers of students in these categories in the population. The standard errors of these estimates and coefficients of variation are also given. These standard errors are estimated by the jackknife technology given in Chapter 7. The coefficient of variation of an estimate is the standard error of the estimate expressed as a percent of the estimate.

Table 4-2

Number of Students by Grade/Age
and Type of Assessment

Assessment Type	Grade 3/Age 9		Grade 7/Age 13		Grade 11/Age 17	
	Total	Sum of Weights	Total	Sum of Weights	Total	Sum of Weights
Spiral[1]	21287	3931992	27668	4007907	39753	4136965
Bridge B--Booklet 4[2]	1994	3151352	2032	3008026	1934	3240017
Bridge B--Booklet 5[2]	2048	3121844	2146	3028806	1934	3252949
Excluded Students--Combined Spiral and Bridge B Samples	*	*	*	*	*	*
Bridge A--Booklet 1[3]	2315	3098639	2075	2937402	**	**
Bridge A--Booklet 2[3]	2361	3104555	2054	2950983	**	**
Bridge A--Booklet 3[3]	2256	3112834	2071	2943837	**	**
Excluded Students--Bridge A Samples	*	*	*	*	**	**

[1] Sample for both age and grade using new age definitions
[2] Sample for age only using new age definitions
[3] Sample for age only using old age definitions
*Not available on Version 1.0 tapes
**No assessment conducted for age 17

Table 4-3

Number of Spiral Students, Grade 3/Age 9

	AGE			
	< 9	= 9	> 9	TOTAL
GRADE < 3				
UNWEIGHTED N	0	2342	0	2342
WEIGHTED N	0	432377	0	432377
STANDARD ERROR	-	19486	-	19486
COEFF. OF VAR.	-	4.51	-	4.51
GRADE = 3				
UNWEIGHTED N	1236	13378	3419	18033
WEIGHTED N	208510	2530844	581923	3321278
STANDARD ERROR	15176	9749	15623	14611
COEFF. OF VAR.	7.28	0.39	2.68	0.44
GRADE > 3				
UNWEIGHTED N	0	912	0	912
WEIGHTED N	0	178337	0	178337
STANDARD ERROR	-	18177	-	18177
COEFF. OF VAR.	-	10.19	-	10.19
GRADE TOTAL				
UNWEIGHTED N	1236	16632	3419	21287
WEIGHTED N	208510	3141559	581923	3931992
STANDARD ERROR	15176	13874	15623	18966
COEFF. OF VAR.	7.28	0.44	2.68	0.48

Table 4-4

Number of Spiral Students, Grade 7/Age 13

	AGE			
	< 13	= 13	> 13	TOTAL
GRADE < 7				
UNWEIGHTED N	0	2775	0	2775
WEIGHTED N	0	646504	0	646504
STANDARD ERROR	-	23559	-	23559
COEFF. OF VAR.	-	3.64	-	3.64
GRADE = 7				
UNWEIGHTED N	1751	16413	5363	23527
WEIGHTED N	257520	2062373	759124	3079017
STANDARD ERROR	16768	3873	16735	10959
COEFF. OF VAR.	6.51	0.19	2.20	0.36
GRADE > 7				
UNWEIGHTED N	0	1366	0	1366
WEIGHTED N	0	282386	0	282386
STANDARD ERROR	-	22763	-	22763
COEFF. OF VAR.	-	8.06	-	8.06
GRADE TOTAL				
UNWEIGHTED N	1751	20554	5363	27668
WEIGHTED N	257520	2991263	759124	4007907
STANDARD ERROR	16768	12922	16735	16317
COEFF. OF VAR.	6.51	0.43	2.20	0.41

Table 4-5

Number of Spiral Students, Grade 11/Age 17

	AGE			
	< 17	= 17	> 17	TOTAL
GRADE < 11				
UNWEIGHTED N	0	5613	0	5613
WEIGHTED N	0	692940	0	692940
STANDARD ERROR	-	16672	-	16672
COEFF. OF VAR.	-	2.41	-	2.41
GRADE = 11				
UNWEIGHTED N	3264	23967	4707	31938
WEIGHTED N	356364	2270022	561849	3188235
STANDARD ERROR	17845	3696	16124	8536
COEFF. OF VAR.	5.01	0.16	2.87	0.27
GRADE > 11				
UNWEIGHTED N	0	2202	0	2202
WEIGHTED N	0	255790	0	255790
STANDARD ERROR	-	17040	-	17040
COEFF. OF VAR.	-	6.66	-	6.66
GRADE TOTAL				
UNWEIGHTED N	3264	31782	4707	39753
WEIGHTED N	356364	3218752	561849	4136965
STANDARD ERROR	17845	8501	16124	13025
COEFF. OF VAR.	5.01	0.26	2.87	0.31

Chapter 5

DATA COLLECTION AND SCORING

Chapter 5: DATA COLLECTION AND SCORING

5.1 INTRODUCTION

In addition to sample selection, Westat, Inc., was responsible for field administration for the 1985-86 NAEP assessment. When data collection was completed, assessment instruments were sent to ETS for processing. This chapter outlines the processes of field administration, data collection, entry and editing of data, and data entry quality control.

5.2 FIELD ADMINISTRATION AND DATA COLLECTION

Based on the 1985-86 sampling design, Westat invited more than 120,000 primary, elementary and secondary school students from more than 1,600 schools in 94 PSUs to participate in the 1985-86 assessment. The assessments in each participating school were coordinated by district supervisors. These supervisors were responsible for conducting sessions in their PSUs and maintaining the security of NAEP materials and the confidentiality of assessment data.

District supervisors provided school personnel with general information about NAEP and worked with them to ensure student attendance at each assessment session. Each supervisor was assisted by two exercise administrators. Supervisors and exercise administrators were responsible for establishing assessment schedules, ensuring that schools were properly prepared for sessions, completing student sample selection in each school, collecting and checking all NAEP materials, and returning materials to ETS.

5.3 DATA ENTRY SYSTEM

The 1985-86 assessment booklets were designed to be read by a computerized scanning device. Respondents filled in ovals next to their response choice(s) for each item. When the instruments were received at ETS, they were processed by the scanning department. The scanned data were transmitted to magnetic tape, then to a computer database.

As the database was created, the data was edited; data in error were indicated on an error listing. A data entry system was developed to correct data discrepancies and to enter data from booklets that were unable to be scanned. The intelligent, direct data entry system allowed entry, verification, and resolution of the data. As each booklet was presented, the system automatically prepared for appropriate exercise sets by setting up corresponding formats or data descriptions. The system also permitted on-line editing of data as they were entered.

5.3.1 Editing

Editing included an assessment of the internal logic and consistency of the data received. For example, data were examined for non-existent school IDs, out-of-range values, and illogical or inconsistent responses. Where possible, conflicts in the data were resolved. If resolution was not possible, the information was left in the form in which it was received.

5.3.2 Multiple Responses

In cases where students, teachers, or school personnel provided more than one response to a single-response background or attitude item, specific guidelines were developed to incorporate these responses consistently and accurately. None of the guidelines applied to achievement items; multiple responses on achievement items were coded as such. The first item in block $9R_1$ is presented below as an example of how multiple responses were resolved.

How often do you do each of the following things when you study for a test?

	Almost every time	More than half the time	Almost half the time	Less than half the time	Never or hardly ever
Read the material over a few times	○	○	○	○	○

The purpose of this item was to find out whether a student ever "read the material over a few times." If the student were to fill in more than one oval and the response data were coded merely "Multiple Response", no information about the student's habits would be recorded. To retain multiple response data for this item, combinations of responses were assigned values. First, a number was assigned to each of the possible responses:

1 = Almost every time
2 = More than half the time
3 = Almost half the time
4 = Less than half the time
5 = Never or hardly ever

Next, these values were combined to form new values:

1 & 2, 1 & 3 = 2
1 & 4, 1 & 5, 2 & 3, 2 & 4 = 3
2 & 5, 3 & 4, 3 & 5 = 4
4 & 5 = 5
All other combinations = 9

The appropriate new values were entered whenever multiple responses were encountered for this item.

5.4 PROFESSIONALLY SCORED ITEMS

As in earlier assessments, the 1985-86 session included items which required open-ended responses. Open-ended response items were administered for reading, science, and mathematics. NAEP recruited and trained a team of readers to score these items. The readers were trained using scoring guides that defined possible score points for each item. Reading responses were scored on a 4-, 5-, or 8-point scale, depending upon the item. Science responses were scored on a 2-, 3-, or 4-point scale, depending upon the item. Mathematics responses were scored as either correct, incorrect or "no response." Reading and science responses were scored by the same readers; a different group of readers scored the mathematics responses.

All readers used scoring guides specifically designed for each item. Some of the items had been used in previous assessments; others were developed for the 1985-86 session. To ensure comparability of results, items from previous assessments were scored using scoring guides from previous assessments. The guides for new items were constructed using data from field testing.

Twenty percent of the reading and science scores were subjected to reliability checks in which the responses were read and scored by a second reader who was not provided with the score assigned by the first. Both scores are retained on the data tapes. Ten percent of the mathematics scores were subjected to a "correctness" check by a second reader, who corrected any mathematics scores found to be in error.

5.5 QUALITY CONTROL

NAEP measured the accuracy of its data entry operation to determine how precisely the data moved from receipt of the instrument to the subsequent machine-readable dataset. For this purpose, a number of student booklets and questionnaires were selected at random and compared, character by character, with their representation on a disk file. The number of booklets and questionnaires involved in quality control checks was based upon the number needed to establish a statistically reassuring conclusion about the accuracy of the entire data entry operation. Student booklets and questionnaires for all three grade/ages were included in quality control checks.

Chapter 6

REPORTING SUBGROUPS AND OTHER VARIABLES

Chapter 6: REPORTING SUBGROUPS AND OTHER VARIABLES[1]

6.1 INTRODUCTION

In addition to reporting national results, NAEP reports information by several student subgroups (results are not reported for individual students). Some subgroup data are taken directly from responses to assessment items; some are derived from responses to two or more different items. This chapter defines the reporting subgroups and explains how their data are derived.

This chapter also explains the data variables on the tapes that have been taken from sources other than the assessment instruments.

6.2 REPORTING SUBGROUPS

NAEP reports information about student subgroups defined by sex, race/ethnicity (both observed and imputed), region of the country, grade, age, level of parent's education, and size and type of community. The derivation of each of these subgroups is explained in the following sections.

6.2.1 Sex (DSEX)

The variable SEX on the student file is the sex of the student being assessed, as taken from school records. For a few students, data for this variable was missing, and was imputed after the assessment by ETS. The resulting variable DSEX on the student file contains a value for every student and should be used for sex subgroup comparisons.

6.2.2 Observed Race/Ethnicity (RACE)

The variable RACE on the student file is the race/ethnicity of the student being assessed, as observed and recorded by the exercise administrator. The observed definition of student race/ethnicity was the only one used in NAEP assessments prior to 1983-84. This variable should be used for race/ethnicity subgroup comparisons to assessments conducted before 1983-84.

---

[1] Data collected from the Teacher Questionnaire and the Excluded Student Questionnaire, and sampling weights for schools, excluded students, and teachers associated with students are not contained on the Version 1.0 data tapes, but will be included in Version 2.0.

6.2.3 Imputed Race/Ethnicity (DRACE)

The variable DRACE on the student file is an imputed definition of race/ethnicity, derived from up to three sources of information. This variable can be used for race/ethnicity subgroup comparisons within the 1985-86 assessment and between the 1985-86 and 1983-84 assessments.

Two common background items were used in the determination of race/ethnicity. The items were included in every spiral and bridge assessment booklet, as follows:

Common Background Item Number 2:

2. If you are Hispanic, what is your Hispanic background?

○ I am not Hispanic
○ Mexican, Mexican American, or Chicano
○ Puerto Rican
○ Cuban
○ Other Spanish or Hispanic background

Students who responded to item number 2 by filling in one of the ovals (○) two through five were considered Hispanic. For students who filled in the first oval, or did not respond to the item, or provided information which was illegible or which could not be classified, responses to item number 1 were examined in an effort to determine race/ethnicity. Item number 1 read as follows:

Common Background Item Number 1:

1. Which best describes you?

○ White
○ Black
○ Hispanic (Mexican, Mexican American, Chicano, Puerto Rican, Cuban, or other Spanish or Hispanic background)
○ Asian or Pacific Islander
○ American Indian or Alaskan Native
○ Other (What?) ________

Students who filled in the first oval were considered White; second oval were considered Black; third oval were considered Hispanic; fourth oval were considered Asian or Pacific Islander; fifth oval were considered American Indian or Alaskan Native. If a student responded by filling in the sixth oval ("Other"), provided illegible information or information which could not be classified, or did not respond at all, observed race/ethnicity (RACE), if provided by the exercise administrator, was used.

Imputed race/ethnicity could not be classified for a student who did not respond to background items 1 or 2, and for whom an observed race/ethnicity was not provided.

Table 6-1 summarizes the procedure used to determine race/ethnicity.

6.2.4 Size and Type of Community (STOC, SSTOC)

NAEP assigned each participating school to one of seven size and type of community categories. These categories are included as the variable STOC on the student file and SSTOC on the school file. The categories were designed to provide information about the communities in which the schools were located.

The STOC reporting categories consist of three "extreme" types of communities and four "residual" community sizes. Schools were placed into STOC categories based upon information about the type of community, the size of its population and upon an occupational profile of residents provided prior to the assessment by school principals. The principals completed estimates of the percentage of students whose parents fit into each of six occupational categories.

Schools in extreme rural and low or high metropolitan areas were ranked in descending order according to the occupational profile, the type of community, and the size of its population. The top 10 percent of these schools were assigned to the extreme STOC categories (1, 2 and 3) below. The remaining schools were classified according to one of the four residual STOC categories. The three extreme STOC categories are as follows:

STOC 1 - Extreme Rural:

This category was used for schools in rural areas where a high proportion of adults was farmers or farm workers and a low proportion of professional, managerial, or factory workers. At least some of the students in these schools were from open country or places with a population of less than 10,000.

STOC 2 - Low Metro:

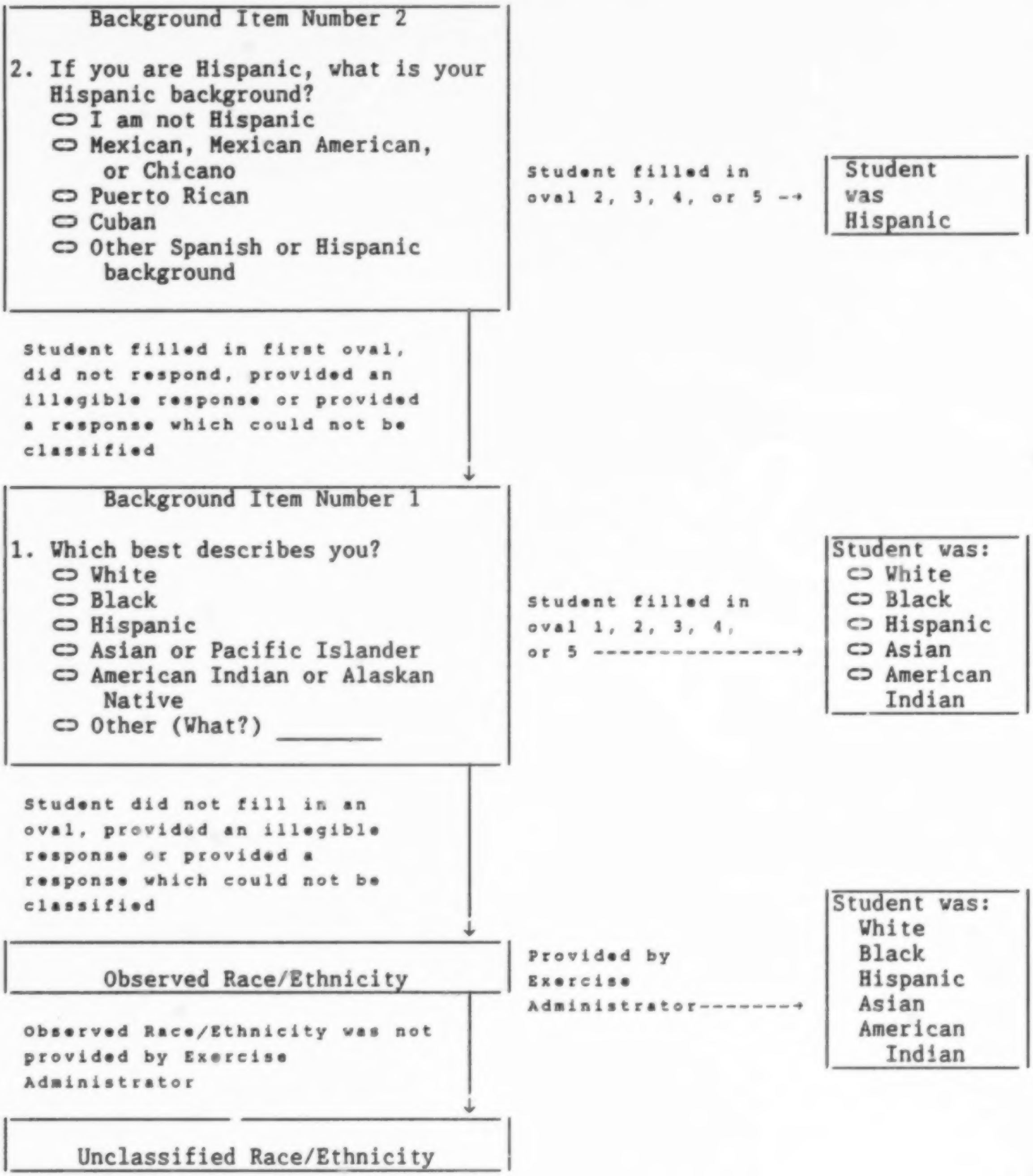
The low metro STOC category was used for schools in areas where a high proportion of the adult population was either not regularly employed or on welfare and a low proportion was employed in professional or managerial positions. The schools in STOC 2 were located in cities, or the urbanized area of cities, with a population greater than 200,000.

STOC 3 - High Metro:

High metro schools were located in city areas where a high proportion of adults was employed in professional or managerial positions and a low proportion factory or farm workers, not regularly employed, or on welfare. STOC 3 schools were located in

Table 6-1

Determining Imputed Race/Ethnicity (DRACE)



Background Item Number 2

2. If you are Hispanic, what is your Hispanic background?
   - ○ I am not Hispanic
   - ○ Mexican, Mexican American, or Chicano
   - ○ Puerto Rican
   - ○ Cuban
   - ○ Other Spanish or Hispanic background

Student filled in oval 2, 3, 4, or 5 → Student was Hispanic

Student filled in first oval, did not respond, provided an illegible response or provided a response which could not be classified ↓

Background Item Number 1

1. Which best describes you?
   - ○ White
   - ○ Black
   - ○ Hispanic
   - ○ Asian or Pacific Islander
   - ○ American Indian or Alaskan Native
   - ○ Other (What?) _______

Student filled in oval 1, 2, 3, 4, or 5 → Student was:
- ○ White
- ○ Black
- ○ Hispanic
- ○ Asian
- ○ American Indian

Student did not fill in an oval, provided an illegible response or provided a response which could not be classified ↓

Observed Race/Ethnicity

Provided by Exercise Administrator → Student was:
- White
- Black
- Hispanic
- Asian
- American Indian

Observed Race/Ethnicity was not provided by Exercise Administrator ↓

Unclassified Race/Ethnicity

cities or the urbanized area of cities with populations greater than 200,000.

Schools which did not fall into STOC 1, 2 or 3 were classified according to four "residual" STOC categories depending upon the size of the community in which they were located. The four residual STOC reporting categories are as follows:

STOC 4 - Main Big City:

STOC 4 schools were located within the limits of cities with populations greater than 200,000 but not classified as High or Low Metro.

STOC 5 - Urban Fringe:

The schools assigned to STOC 5 were located in the urbanized area, but outside the limits, of cities with populations over 200,000, but not classified as Low or High Metro.

STOC 6 - Medium City:

STOC 6 schools were located in cities with populations of between 25,000 and 200,000 which did not classify as fringe areas for big cities.

STOC 7 - Small Place:

The schools assigned to STOC 7 were located in communities with populations of less than 25,000. These communities were not located in the urbanized areas of big cities and could not be classified as Extreme Rural.

6.2.5 Parental Education (PARED)

The variable PARED on the student file is derived from responses to two common background items (B003501 and B003601). Students were asked to indicate the extent of their father's education (item B003601) by choosing one of the following:

○ He did not finish high school;
○ He graduated from high school;
○ He had some education after high school;
○ He graduated from college; or
○ I Don't Know.

Students were asked to provide the same information about the extent of their mother's education (item B003501) by choosing one of the following:

○ She did not finish high school;
○ She graduated from high school;
○ She had some education after high school;
○ She graduated from college; or
○ I Don't Know.

The information was combined into one parental education reporting category (PARED), as follows:

If a student indicated the extent of education for only one parent, that level was included in the data. If a student indicated the extent of education for both parents, the higher of the two levels was included in the data. If a student indicated that he or she did not know the level of education for both parents or indicated that he or she did not know the level of education for one parent and did not respond for the other, the parental education level was classified as unknown. If the student did not respond for both parents, the student was recorded as having provided no response.

6.2.6 Region (REGION, SREGION)

In addition to overall responses, NAEP computed data for four geographical regions in the United States. These data are retained on the data tapes as the variable REGION on the student file and SREGION on the school file. Table 6-2 outlines the assignment of individual states to each region. (Note: Alaska and Hawaii were not included in the sample.)

Table 6-2
Geographic Regions

NORTHEAST:		SOUTHEAST:	
Connecticut	New Hampshire	Alabama	Mississippi
Delaware	New Jersey	Arkansas	North Carolina
District of Columbia	New York	Florida	South Carolina
	Pennsylvania	Georgia	Tennessee
Maine	Rhode Island	Kentucky	Virginia
Maryland	Vermont	Louisiana	West Virginia
Massachusetts			

CENTRAL:		WEST:	
Illinois	Missouri	Alaska	New Mexico
Indiana	Nebraska	Arizona	Oklahoma
Iowa	North Dakota	California	Oregon
Kansas	Ohio	Colorado	Texas
Michigan	South Dakota	Hawaii	Utah
Minnesota	Wisconsin	Idaho	Washington
		Montana	Wyoming
		Nevada	

6.2.7 Grade (DGRADE. MODGRD)

To enhance the usefulness of 1985-86 assessment data, NAEP sampled students by grade as well as by age. Students were also sampled in 1983-84 by both grade and age; the student cohorts in that assessment were grade 4/age 9, grade 8/age 13, and grade 11/age 17. The ages sampled matched those sampled in earlier assessments; each grade sampled represented the corresponding modal grade, or the grade attended by most students of the particular age.

In 1985-86, with changes in age definition and time of testing for ages 9 and 13, the corresponding modal grades also changed, from grades 4 and 8 to grades 3 and 7. Age definition and time of testing did not change for the older student cohort, which remained grade 11/age 17.

The main (spiral) sample includes many students in each cohort who were both age-eligible (age 9, 13, or 17) and grade-eligible (attending respectively grade 3, 7, or 11). However, because NAEP collected data by grade or age, each cohort also includes students who were age-eligible but not in the modal grade, and students who were grade-eligible but not of the modal age (the age of most students attending the particular grade).

The bridge studies performed to assess the effects of the changes in age definition and time of testing (Bridge A) and mode of administration (Bridge B) sampled students by age only.

Results for students in a particular grade can be selected using the variable DGRADE, the student's actual grade at time of testing, on the student file.

Results for students in three ranges of grades (i.e., students in grades below the modal grade, at the modal grade, and above the modal grade) can be selected using the variable MODGRD on the student file.

Table 6-3 compares MODGRD values with their corresponding DGRADE values for the three grade/ages.

6.2.8 Age (DAGE, MODAGE)

For the 1985-86 main (spiral) sample, student age for all three grade/age cohorts was calculated as of September 30, 1986. Because NAEP collected data by grade or age, each spiral sample student cohort includes students who were both age-eligible and grade-eligible, students who were age-eligible but not in the modal grade, and students who were grade-eligible but not of the modal age. The modal ages for each grade/age cohort in the spiral sample were defined by the following birthdates, based on the school year:

Age 9: born between October 1, 1976 and September 30, 1977
Age 13: born between October 1, 1972 and September 30, 1973
Age 17: born between October 1, 1968 and September 30, 1969

Table 6-3

Comparison of MODGRD and DGRADE Values

	DGRADE Grade 3/Age 9		DGRADE Grade 7/Age 13		DGRADE Grade 11/Age 17	
MODGRD	Spiral & Bridge B	Bridge A	Spiral & Bridge B	Bridge A	Spiral & Bridge B	Bridge A
1 (< modal grade)	1 2	1 2 3	4 5 6	5 6 7	7 8 9 10	--
2 (= modal grade)	3	4	7	8	11	--
3 (> modal grade)	4	5	8 9	9	12	--

Table 6-4

Comparison of MODAGE and DAGE Values

MODAGE	DAGE - Age 9	DAGE - Age 13	DAGE - Age 17
1 (< modal age)	7 8	9 10 11 12	14 15 16
2 (= modal age)	9	13	17
3 (> modal age)	10 11 12 13 14	14 15 16 17 18	18 19 20 21 22 23

Because the definitions of ages 9 and 13 were changed in the 1985-86 assessment, a bridge study (Bridge A) was conducted to determine the effects of the change. Age eligibility for the Bridge A samples was based on the previous definition of age (by calendar year) and was calculated as of December 31, 1986.

The Bridge A age 9 samples include only students born between January 1 and December 31, 1976. Students with those birthdates who were in a grade lower than second grade or higher than fifth grade were not included in the Bridge A samples.

The Bridge A age 13 samples include only students born between January 1 and December 31, 1972. Students with those birthdates who were in a grade lower than sixth grade or higher than ninth grade were not included in the Bridge A samples.

A second bridge study, Bridge B, was conducted to determine the effects of a change in mode of administration. Age eligibility for the Bridge B samples was based on the new definition of age (by school year) and was calculated as of September 30, 1986.

The Bridge B samples include only students born between October 1, 1976 and September 30, 1977 (age 9); students born between October 1, 1972 and September 30, 1973 (age 13); and students born between October 1, 1968 and September 30, 1969 (age 17).

Results for students in a particular age can be selected using the variable DAGE, the student's actual age at time of testing, on the student file.

Results for students in three ranges of ages (i.e., students in ages below the modal age, at the modal age, or above the modal age) can be selected using the variable MODAGE on the student file.

Table 6-4 compares MODAGE values with their corresponding DAGE values for the three student age cohorts.

6.3 OTHER VARIABLES

Several variables on the data tapes have been taken from sources other than the student assessment instruments or assessment survey questionnaires. These are described in the following sections.

6.3.1 Principal Questionnaire Variables

Prior to the assessment, Westat, Inc., distributed a questionnaire to each school principal to gather data about school characteristics, including school enrollment and attendance, parents' occupations, and student ethnicity. These data variables are retained on the school file, and are identified in the data layouts by "(PQ)" in the SHORT LABEL field.

The NAEP variables

SPCTAMI (percent American Indian)
SPCTASI (percent Asian/Pacific Islander)
SPCTHSP (percent Hispanic)
SPCTBLK (percent Black)
SPCTWHT (percent White)

represent percents of student ethnicity in each school. The data from these variables were derived from the percent-ethnicity variables from the Principal Questionnaire: PPCTAMI, PPCTASI, PPCTHSP, PPCTBLK, and PPCTWHT. Missing data for these variables were supplemented with corresponding data from another source, Quality Education Data, Inc. (see below). If percent-ethnicity data was not available from either source, the derived variables indicate "Missing" for those schools.

6.3.2 Quality Education Data (QED) Variables

The data tapes contain several variables obtained from information supplied by Quality Education Data, Inc. (QED). QED maintains and updates annually lists of schools showing grade span, total enrollment, school district, principal's name and other information for each school. Some of this information has been used for the following variables:

Student Files			School Files		
ORSHPT	IDP	SPECED	SORSHPT	SIDP	SSPECED
NTEACHC	CAI	LIB/MED	SNTCHC	SCAI	SLIBMED
SCHTYPE	URBAN	INDARTS	SSCHTYP	SURBAN	SINDART
GRSPAN	NSTUDA	ADULTED	SGRSPAN	SNSTUD	SADLTED
ENROLL	NTEACHA	PCENROL	SENROLL	SNTEACH	SPCENROL

Most of the QED variables are sufficiently defined in the data codebooks. Explanations of others are provided below.

ORSHPT and SORSHPT are the Orshansky Percentile, an indicator of relative wealth which specifies the percentage of district school-age children who fall below the poverty line.

IDP and SIDP represent, at the school district level, dollars per student spent for textbooks and supplemental materials.

ADULTED and SADLTED indicate whether or not adult education courses are offered at the school site.

URBAN and SURBAN define the school's urbanicity: Urban (central city), Surburban (area surrounding central city, but still located within the counties constituting the Standard Metropolitan Statistical Area; or Rural (area outside any Standard Metropolitan Statistical Area).

PCENROL and SPCENROL indicate the percent increase or decrease in school enrollment from 1980.

Chapter 7

CONDUCTING STATISTICAL ANALYSES OF 1985-86 NAEP DATA

Chapter 7: CONDUCTING STATISTICAL ANALYSES OF 1985-86 NAEP DATA[1]

7.1 INTRODUCTION

Users of NAEP data should be aware of the special properties of the NAEP database that affect the validity of conventional techniques of statistical inference. Because a complex sampling scheme, rather than simple random sampling, was used to collect NAEP data, standard procedures for statistical inference should not be applied to the NAEP data without modification.

In the NAEP sampling scheme, students do not have an equal probability of being selected. Therefore, as in all complex surveys, each student has been assigned a sampling weight. The larger the probability of selection for students within a particular demographic group, the smaller the weights for those students will be. When computing descriptive statistics or conducting inferential procedures, one should weight the data for each student. Performance of statistical analyses without weights can lead to misleading results.

Another way in which the complex sample design used by NAEP differs from simple random sampling is that the NAEP sampling scheme involves the selection of clusters of students who come from the same school, as well as clusters of schools that come from the same geographical region. As a result, observations are not independent of one another as they are in a simple random sample. Therefore, use of ordinary formulas for estimating the standard error of sample statistics, such as means, proportions or regression coefficients, will result in values that are too small. The standard error, which is a measure of the variability of a sample statistic, gives an indication of how well that statistic estimates the corresponding population value. It is used in conducting tests of statistical significance. If conventional simple random sampling (SRS) formulas are used to compute standard errors, "too many" statistically significant results will occur.

Yet another effect of the NAEP sampling scheme is a reduction of the effective degrees of freedom. In a simple random sample, the degrees of freedom of a variance estimate are based on the number of subjects. In the NAEP design, the degrees of freedom are a function of the number of clusters of primary sampling units (PSUs) and the number of strata in the design, rather than the number of subjects (see Chapter 4 for a discussion of the sample design). Therefore, the ordinary formulas for obtaining degrees of freedom are not valid with the NAEP data.

---

[1] Data collected from the Teacher Questionnaire and the Excluded Student Questionnaire, and sampling weights for schools, excluded students, and teachers associated with students are not contained on the Version 1.0 data tapes, but will be included in Version 2.0.

The following sections will outline the procedures used in NAEP to account for these special properties of the NAEP data. Section 7.2 discusses the use of weights to account for the differential sampling rates. Section 7.3 discusses jackknife procedures that can be used to estimate sampling variability. Section 7.4 provides an approximate procedure, using design effects, for estimating sampling variability. Although this technique is less precise than the jackknife, it involves considerably less computation. It is expected that the resultant degree of accuracy will be acceptable to most users of the NAEP data.

7.2 USING WEIGHTS TO ACCOUNT FOR DIFFERENTIAL REPRESENTATION

The 1985-86 NAEP used a complex sample design to obtain the students who were assessed. The goal of this design was to obtain a series of samples (for the various ages and grades) from which estimates of population and subpopulation characteristics could be obtained with reasonably high precision (as measured by low sampling variability). To accomplish this goal, NAEP used a multistage cluster sample design (described in Chapter 4) in which the probabilities of selection of the clusters was proportional to measures of their size. To ensure adequate precision in the estimation of the characteristics of various subpopulations of interest, some subpopulations (corresponding to students from areas with high concentrations of Black, Hispanic and/or Asian students) were deliberately sampled at approximately twice the normal rate to obtain larger samples of respondents from those subpopulations. The result of these differential probabilities of selection is a series of achieved samples, each containing proportionately more members of certain subgroups than there are in the population.

Appropriate estimation of population characteristics must take this disproportional representation into account. This is accomplished by assigning a weight to each respondent, where the weights properly account for the sample design and reflect the appropriate proportional representation of the various types of individuals in the population. These weights also include adjustments for nonresponse and adjustments (known as post-stratification adjustments) designed to make sample estimates of certain subpopulation totals conform to external, more accurate, estimates. Details about the computation of these weights appear in Chapter 4. For the present purpose, it is sufficient to note that these weights should be used for all analyses, whether exploratory or confirmatory.

The 1985-86 database includes a number of different samples from several populations. Each of these samples has its own set of weights which are to be used to produce estimates about the characteristics of the population addressed by the sample (the target population). The various samples, their target populations, and their weights are discussed in the following sections.

7.2.1 The Spiral Samples of Students

These samples, one for each of the three grade/age combinations, consist of all students assessed in the main (spiral) assessment. The target population for each of these samples consists of all students who are in the specified grade/age combination who were deemed assessable by their school. The weight to be used for all analyses at the student level, based on these samples, is the student full-sample weight (WEIGHT).

The sum of weights for all students in a particular grade/age spiral sample estimates the total number of assessable students in the population who are of that grade/age combination. In most cases such an estimate is of minor interest; a researcher is more likely interested in the number of students at either a grade or an age. An estimate of the number of assessable students at an age level can be obtained by summing the weights of only the age-eligible students; and an estimate of the number of assessable students in a given grade comes by summing the weights of only the grade-eligible students. In general, an estimate of the number of assessable students in the population who possess some characteristic, such as third grade students in the Northeast, comes by summing the weights of all respondents to the spiral assessment who have that characteristic.

An estimate of the proportion of students in the population who possess some characteristic is obtained as the ratio of the estimated total number of students with that characteristic to the estimated population total. This estimate can be restricted to subpopulations. For example, the estimated proportion of all assessable third grade students who are also in the northeast is

$$\frac{\text{WTOT (3rd grade and NE)}}{\text{WTOT (3rd grade)}}$$

where WTOT (3rd grade and NE) is the sum of the weights of all students in the spiral sample who are both in the third grade and the Northeast and where WTOT (3rd grade) is the sum of the weights of all students in the spiral sample who are in the third grade and from any region.

It is also clearly of interest to estimate the relative proportion of a population (say 3rd grade students) who could correctly respond to an assessment exercise. This proportion is estimated by the ratio

$$P = \frac{\text{WTOT (3rd grade, answered item correctly)}}{\text{WTOT (3rd grade, presented the item)}}$$

where the numerator is the sum of weights of all assessed students who are in the third grade and who correctly responded to the item. Please note that the denominator is the sum of weights of all students who

1) are in the third grade and
2) were presented the item.

This total is less than WTOT (3rd grade) because not all students are presented every item (due to the spiral design). However, the sample of assessed students who were presented the item is itself a representative sample of the entire population of assessable students.

7.2.2 The Bridge B Samples of Students

For each of the three ages (9, 13 and 17--using the new age definitions) there are two bridge samples (from booklets 4 and 5) designed to allow the measurement of the effect of changing from tape-recorded to printed mode of administration. The target population for each of these samples consists of all students who are of the specified age and who would be assessable. The weight to be used for all analyses at the student level, based on any one of these samples, is the student full-sample weight (WEIGHT).

The sum of weights of all respondents to a particular bridge sample estimates the total number of age-eligible students in the country who would be assessable. Since there are two such samples, there are two estimated totals which differ only because of sampling variability. Another estimate can be obtained as the average of these two estimates.

Estimates of totals for subpopulations as well as estimates of proportions can be obtained in the same manner as in the above discussion for the spiral sample.

7.2.3 The Excluded Students from the Spiral and Bridge B Assessments

The excluded students from the combined spiral and Bridge B assessments are a sample from the population of all grade/age eligible students who, in the opinion of their school, would not be assessable. The weight to be used for all analyses of these data, such as for analysis of the responses to the Excluded Student Questionnaire, is the excluded student final weight (XWEIGHT).

Besides conducting weighted analyses on these data to estimate characteristics of the population of excluded students, analogous to the analyses for the assessed students, it is possible to combine the various samples together.

For example, the sum of the weights of all spiral assessed students of a given grade/age combination plus the equivalent sum of weights for the excluded students estimates the total number of in-school students (assessable or not) in the grade/age combination.

7.2.4 The Bridge A Samples of Students

For each of the ages 9 and 13, there are three bridge samples (from booklets 1, 2 and 3) designed to allow the measurement of the effect of changing the age definitions and the time of year the assessment data were collected. Consequently, the target population for each of these samples, while being the assessable age eligibles, is different from the target population of age eligibles in the spiral and Bridge B assessments, the difference being the change in definition of age between the two types of samples.

The weight to be used for all analyses of these data is the student-full sample weight (WEIGHT).

7.2.5 The Excluded Students from the Bridge A Assessments

These are a sample from the population of all age eligible students, using the old age definitions, who would be unassessable. The weight to use for all analyses of these data is the excluded student final weight (XWEIGHT). These data can be combined with the data from any one of the fall/winter bridge samples to estimate characteristics of the total population of in-school students (assessable or not) of the specified age class.

7.2.6 Other Weights

The final teacher-student weights (TSTUWTF) are appropriate for analyses of the Teacher Questionnaire data, where the teacher's characteristics are linked to the students. These will be available for the spiral assessment only.

The teacher-student weights are appropriate for use in estimating the number or percent of students in the total population who have various characteristics; they are not appropriate for use in estimating the number of teachers, or the number of teachers with various characteristics. They are supplied only for a subsample of students who were selected from the full NAEP sample and are appropriate for use in analyses involving the teacher-characteristics of students. For example, these weights would be used to estimate the proportion of students who have Hispanic language arts teachers.

The school weights will be used to analyze, at the school level, the School Questionnaire data.

7.3 PROCEDURES USED BY NAEP TO ESTIMATE VARIABILITY (Jackknifing)

This section describes how the sampling variability of statistics based on the NAEP data can be estimated. This estimate represents the most precise estimate obtainable, given the resources available to NAEP. Better estimators of uncertainty could be obtained with more resources. Because most secondary users will have fewer resources than those available for the NAEP reports, Section 7.4 provides less expensive approximations to the variance estimator described below.

A major source of uncertainty in the estimation of the value in the population of a variable of interest exists because information about the variable is obtained only on a sample from the population. To reflect this fact, it is important to attach to any statistic (e.g., a mean) an estimate of the sampling variability to be expected for that statistic.

Estimates of sampling variability are designed to provide information about how much the value of a given statistic would be likely to change if the statistic had been based on another, equivalent, sample of individuals drawn in exactly the same manner as the achieved sample. Consequently, the estimation of the sampling variability of any statistic must take into account the design of the sample.

The NAEP sample is obtained via a stratified multistage probability sampling design which includes provisions for sampling certain subpopulations at higher rates. Additional characteristics of the sample include adjustments for both nonresponse and post-stratification. The resulting sample has different statistical characteristics from those of a simple random sample. In particular, because of the effects of cluster selection (students within schools, schools within PSUs) and because of effects of nonresponse and post-stratification adjustments, observations made on different students cannot be assumed to be independent of each other (and are, in fact, generally positively correlated). Furthermore, to account for the differential probabilities of selection (and the various adjustments), each student has an associated sampling weight, which must be used in the computation of any statistic and which is itself subject to sampling variability.

Treatment of the data as a simple random sample, with disregard for the special characteristics of the NAEP sample design, will produce underestimates of the true sampling variability.

A procedure known as jackknifing is suitable for estimating sampling errors from such a complex design. This procedure has a number of properties which make it particularly suited for the analysis of NAEP data:

1) The jackknife procedure properly estimates the sampling error arising from the complex sample selection procedure for linear estimates such as simple totals and means, and does so approximately for more complex estimates.

2) It reflects the component of sampling error introduced by the use of weighting factors which are dependent upon the sample data actually obtained.

3) Jackknifing can be adapted readily to the estimation of sampling errors for parameters estimated using statistical modeling procedures, as well as for tabulation estimates.

4) Once appropriate weights are derived and attached to each record, the procedure is straightforward to use for estimating sampling errors. A single set of replicate weights is required for all tabulations and model parameter estimates which may be needed.

The method of applying the jackknife procedure involves first defining pairs of groups of first-stage sampling units. For the 1985-86 NAEP spiral assessment, Westat has defined 76 groups of such first-stage units, paired so that the populations represented by each member of the pair are similar. In many cases, a group consists of a single PSU. This results in 38 pairs of first-stage units. Similar pairings have been defined for the bridge assessments. For the Bridge A assessment there are 33 pairs; there are also 33 pairs for the Bridge B assessment (but these correspond to somewhat different clusters of PSUs). These pairings are identified by the variable JKPAIR on the data tapes; membership within the pair is identified by the variable JKREPL.

The component of the sampling variability attributable to a pair of first-stage units is estimated as the squared difference between the value of the statistic for the complete sample and a pseudoreplicate formed by recomputing the statistic on a specially constructed pseudodataset. This pseudodataset is created from the original dataset by eliminating one member of the pair and replacing it with a copy of the second set of first-stage units in the pair. For computational purposes, the pseudodataset associated with a given pair is the original dataset with a different set of weights, referred to as the student replicate weights SRWT01 through SRWT38 on the data tapes, where SRWTi is for the i^{th} pair. This set of weights allows measurement of the total effect of replacing one member of the pair with a copy of the other, including adjustments for nonresponse and post-stratification. The pseudoreplicate associated with the i^{th} pair for a given statistic is obtained by recalculating the statistic using the weights SRWTi instead of the sampling weights.

The student replicate weight, SRWTi, for the i^{th} pair of first-stage units was computed as follows:

Let W_B be the base weight of a student, where the base weight accounts for the probabilities of selection but does not include nonresponse or post-stratification adjustments.

Then $\text{SRWTi} = f_i^{NR} f_i^{PS} W_{Bi}$

where

$$W_{Bi} = \begin{cases} 0 & \text{JKPAIR} = i,\ \text{JKREPL} = 2 \\ \text{JKFAC} \cdot W_B & \text{JKPAIR} = i,\ \text{JKREPL} = 1 \\ W_B & \text{JKPAIR} \neq i \end{cases}$$

is the replicate base weight formed by replacing the second member of the pair by the first, JKFAC is a constant multiplier (usually equal to 2) designed to maintain certain population totals, and where f_i^{NR} and f_i^{PS} are, respectively, nonresponse and post-stratification adjustment factors based on these replicated base weights.

As a specific example of the use of the student replicate weights, let $t(\underline{y},\underline{w})$ be any statistic which is a function of the sample responses $\underline{y}$ and the weights $\underline{w}$ and which estimates population value T. For example, t could be a weighted mean, a weighted percent-correct point or a weighted regression coefficient. The $t(\underline{y},\underline{w})$, computed with the sampling weights (WEIGHT on the data tapes) is the appropriate sample estimate of T. To compute $\hat{\mathrm{Var}}(t)$, the sampling variance for this statistic, proceed in the following manner:

1) For each of the 38 pairs of first-stage units, compute the associated pseudo-replicate for the statistic. For the i^{th} pair, this is

$$t_i = t(\underline{y},\underline{\mathrm{SRWTi}})$$

which is the statistic t recalculated by using SRWTi instead of the sampling weights.

2) The sample variance of t is

$$\hat{\mathrm{Var}}(t) = \sum_{i=1}^{38} (t_i - t)^2 \quad .$$

This estimation technique is called the multi-weight jackknife approach. Tables 9-7 and 9-8 in Chapter 9 provide SPSS-X and SAS code for carrying out the above in the special case of a weighted mean.

A similar procedure is followed to estimate the sampling variability for statistics based on any of the bridge samples. The only difference is in the number of pairs (and hence replicate weights) used.

Replicate weights have been provided for the following samples:

1) Spiral students	SRWT01 to SRWT38
2) Bridge B students	SRWT01 to SRWT33
3) Bridge B excluded students	EXWT01 to EXWT38
4) Bridge A students	SRWT01 to SRWT33
5) Bridge A excluded students	EXWT01 to EXWT33
6) Teacher-student sample	TSWT01 to TSWT38

7.3.1 Degrees of Freedom of the Jackknifed Variance Estimate

Note that the jackknife procedure estimate the sampling variability of the statistic by assessing the effect of change in the sample at the level of clusters of first-stage units. For this reason, the number of degrees of freedom of the variance estimate Vâr(t) will be at most equal to the number of pairs. The number of degrees of freedom, which indicates the variability of the variance estimate, equals the number of independent pieces of information used to generate the variance. In the current case, for the spiral sample, the pieces of information are the 38 squared differences $(t_i - t)^2$, each supplying at most one degree of freedom, regardless of how many individuals were sampled within any PSU. (There are fewer pairs with the bridge samples.)

Increasing the number of individuals sampled within any PSU results in a lower estimate of sampling variability because the within-PSU component is reduced. This, however, does not improve the estimation of the between-PSU component of variability, which depends on the number of PSUs selected.

The number of degrees of freedom of the sample variance estimator can be less than the number of pairs (38) when a few of the squared differences $(t_i - t)^2$ are markedly different in magnitude than the remainder. An extreme case of this is when one or more of the t_i are identical to t, so that $(t_i - t)^2 = 0$. This may happen, for example, when the statistic t is a mean for a subgroup and no members of that subgroup come from the pair i. Such a pair contributes zero degrees of freedom to the variance.

An estimate of the effective number of degrees of freedom for Vâr(t) comes from an approximation due to Satterthwaite (1946). (See Cochran, 1977, p. 96, for a discussion.)

The effective number of degrees of freedom using this approximation is

$$df_{eff} = \frac{\left(\sum_{i=1}^{38} (t_i - t)^2 \right)^2}{\sum_{i=1}^{38} (t_i - t)^4} ,$$

which is never larger than 38.

7.4 APPROXIMATIONS

The major computational load in computing uncertainty measures for any statistic exists in the computation of the uncertainty due to sampling variability. As mentioned earlier, the procedure detailed in Section 7.3.1 requires that the statistic be recomputed 39 times to obtain an estimate of the sampling variance of the statistic. This section describes how to approximate the sampling variability for any statistic.

As indicated in Section 7.3.1, it is inappropriate to estimate the sampling variability of any statistic based on the NAEP database by using simple random sampling (SRS) formulas. These formulas, which are the ones used by most standard statistical software such as SPSS and SAS, will produce variance estimates which are generally much smaller than is warranted by the sample design.

It may be possible to account approximately for the effects of the sample design by using an inflation factor, the design effect, developed by Kish (1967) and extended by Kish and Frankel (1974). The design effect for a statistic is the ratio of the actual variance of the statistic (taking the sample design into account) over the conventional variance estimate based on the same number of elements. To avoid sources of bias due to improper representation, this conventional estimate must use the sampling weights. The design effect may be used to adjust error estimates based on simple random sampling assumptions to account approximately for the effect of the design. In practice, this is often accomplished by dividing the total sample size by the design effect and using this effective sample size in the computation of errors. Note that the value of the design effect depends on the type of statistic computed and the variables considered in a particular analysis as well as the clustering effects occurring among sampled elements.

Based on empirical results and theoretic considerations, Kish and Frankel (1974) have developed several conjectures about design effects:

1) Generally, the design effects for complex statistics from complex samples are greater than 1, causing variances based on simple random sampling assumptions to tend to be underestimates.

2) The design effects for complex statistics (such as regression coefficients) tend to be smaller than the corresponding design effects for means of the same variables. Hence, these latter estimates, which are more easily computed, tend to give overestimates of the design effects of complex statistics.

3) The design effects of complex statistics tend to resemble those of means; variables with a high design effect of the mean also tend to have high design effects for complex statistics involving those variables.

To incorporate the design effect idea in a statistical analysis, proceed in the following manner:

1) For a given class of statistics (e.g., means, percentile points, regression coefficients), compute the jackknife variance as in Section 7.3.1 for a number of cases. The cases should cover the range of situations for which the approximation is to be used. If various subpopulations are to be considered, it is important to have information on the relative variability within each subgroup. This is especially important if certain subgroups are more highly clustered in the sample.

2) For the identical cases, compute the conventional estimate of the variance. This estimate must take the sample weights into account to avoid problems of bias due to improper representation. To account properly for the difference between the number of individuals being sampled and the total of the sampling weights, the weights should be scaled so that their sum equals the sample size.

3) For each case, compute the design effect where the design effect for case j is

$$\mathrm{deff}_j = \mathrm{Var}_{JK}(t_j)/\mathrm{Var}_{CON}(t_j) \quad ,$$

the ratio of the jackknife variance estimate of the statistic to its conventional variance estimate.

4) If the design effects for the various cases are tolerably similar, choose an overall composite design effect. If the design effects for certain subgroups appear to cluster around a markedly different value from the remaining cases, treat those subgroups separately.

5) In the case that a consistent overall design effect has been found:

   a) rescale the weight of each individual so that the sum of the scaled weights is equal to the effective sample size

   $$N_{eff} = \frac{\text{sample size}}{\text{design effect}}$$

   b) conduct a traditional weighted analysis using these scaled weights

6) The degrees of freedom for any variance estimates obtained by using this approach is still at best 38, the number of pairs, as it was for the jackknife. Accordingly, tests of significance produced by standard programs (which will use the effective sample size minus the number of parameters for error degrees of freedom) should be interpreted with extreme caution because they are likely to be too liberal. Significance and inferential procedures are properly based on the smaller error degrees of freedom.

7.5 A NOTE CONCERNING MULTIPLE COMPARISONS

In performing multiple hypothesis tests it is important to consider the fact that if J tests are performed, each with a Type I error rate (the probability of rejecting the null hypothesis when the null hypothesis is true) of α, the Type I error rate for the entire set of contrasts could be as high as $J\alpha$. Therefore, it is desirable to use a multiple comparison procedure to control the overall error rate for the entire set of hypothesis tests. In the present case, it is advantageous to use a procedure that allows control of the error rate for sets of varying size that may include both pairwise and complex comparisons. (An example of a complex contrast is a comparison of one group to the average of two other groups.) The Dunn-Bonferroni approach is, therefore, a good choice. To apply this method in its simplest form, we need only decide at what level we wish to control the setwise error rate (α_s) and then set the Type I error rate for each comparison equal to $\alpha_c = \alpha_s/J$, where J is the number of comparisons.

For example, suppose we wanted to perform three pairwise comparisons between regional groups, as well as one complex comparison, controlling α_s at .05. The Type I error rate for each comparison should be set at $\alpha_c = \alpha_s/J = .05/4 = .0125$. The required critical value can be obtained from a table of the Bonferroni t-statistic (Miller, 1981, p. 238) with the appropriate degrees of freedom.

References

Cochran, W. G. (1977). Sampling techniques (3rd ed.). New York: John Wiley & Sons.

Kish, L. (1967). Survey sampling. New York: John Wiley & Sons.

Kish, L., & Frankel, M. R. (1974). Inference from complex samples. Journal of the Royal Statistical Society, Series B, 36, 1-22.

Miller, R. G. (1981). Simultaneous statistical inference (2nd ed.). New York: Springer.

Satterthwaite, F. E. (1946). An approximate distribution for estimates of variance components. Biometrics, 2, 110-114.

Chapter 8

CONTENT AND FORMAT OF THE DATA TAPES

Chapter 8: CONTENT AND FORMAT OF THE DATA TAPES[1]

8.1 INTRODUCTION

This chapter provides detailed information concerning the contents and format of each data file on the public-use data tapes and of the printed layouts and codebooks produced for the 1985-86 assessment.

8.2 RAW DATA FILES

The raw data files contain one record per student, excluded student, teacher, and school. All raw data files are rectangular, so that record lengths are fixed and a given variable always occurs in the same position on every record within a file. The data files for each grade/age are described in Tables 8-4, 8-5, and 8-6.

Due to the spiral design (see Chapter 3), students were administered different blocks of items during the 1985-86 assessment. As a result, each student record contains blank spaces for the exercise blocks that were not included in his or her assessment booklet (missing by design). Fields are also blank for items which did not appear in booklets due to a printing error (e.g., incorrect block in book, missing pages) and for the professionally scored items that were not included in reliability checks (see Section 5.4 in Chapter 5).

Responses classified as off-task, illegible, omitted or not reached, I Don't Know, or out-of-range have been included in the raw data files, codebooks, and machine-readable catalogs and are coded as indicated in Table 8-1.

8.3 LAYOUTS

With the exception of field types, all information in the layouts has been numerically coded. The layouts include the following information for each data field:

SEQ. NO.	Sequence number. Fields are numbered sequentially to represent the order in which they appear on the data record.

---

[1] Data collected from the Teacher Questionnaire and the Excluded Student Questionnaire, sampling weights for schools, excluded students, and teachers, and IRT parameters for student responses are not contained on the Version 1.0 data tapes. They will be included in Version 2.0.

Table 8-1

Code Definitions

CODE FOR FIELD WIDTH = 1	CODE FOR FIELD WIDTH = 2	CODE DEFINITION
5	55	ILLEGIBLE (items N003104 and N003105 only)
6	66	OFF TASK (items N003104 and N003105 only)
7	77	I DON'T KNOW (items N003104, N003105 and all multiple-choice items)
		NON-RATEABLE (professionally scored items)
8	88	OMITTED or NOT REACHED (all items)
9	99	OUT-OF-RANGE RESPONSE (all items)

FIELD NAME	A short name (of up to eight characters) which identifies the field. The format of the field names associated with response data follows:
Position 1	identifies nature/source of the response data: B = Common background item within common background block S = Background and attitude item within cognitive block (with two exceptions: item S002701 appears in the common background block for all grade/ages; item S004001 appears in the common background block for grade 7/age 13 and grade 11/age 17) N = Cognitive item within cognitive block C = School questionnaire item T = Teacher questionnaire item X = Excluded student questionnaire item H = U. S. history item L = Literature item
Positions 2 through 5	identify an exercise (Student files) or question (School, Teacher, Excluded Student files). If position 1 is S or N, a zero in position 2 signifies a reading item; a 2 signifies a mathematics item; 4 signifies a science item; and 6 signifies a computer item.
Positions 6 and 7	identify a part within an exercise (Student files) or a part within a question (School, Teacher, Excluded Student files)
Position 8	identifies the block containing an item (Student files only). To save file space, the numeric designation (1 through 11) has been replaced by an alphabetic one (A through K). This position is left blank for questionnaire variables.
COL. POS.	Column position. The start location of the field on the data record.
FIELD WIDTH	The number of characters in the field.
DECIMAL PLACES	The number of digits to the right of the decimal point in the field. The raw data contain implicit decimal points.
TYPE	The files include three field types, as follows:
Type C	(Continuous) fields designate continuous numerical data without fixed ranges.
Type D	(Discrete) fields designate discrete data with a fixed number of responses. Type D fields may include raw item responses or imputed (derived) categorical variables.

Type DI	(Discrete with "I Don't Know") fields designate discrete data with a special code for an "I Don't Know" response. Depending on the field width, the "I Don't Know" value is either 7 or 77.
RANGE	The range of values or of valid responses for a field.
KEY VALUE	The correct response for a multiple-choice item. For some professionally scored items, the key is expressed as a range of values, denoting cut-point scoring for item analysis. (See CORRECT RESPONSE in Section 8.5.)
SHORT LABEL	A brief description of the information in the field.

8.4 CODEBOOKS

Each entry in the codebooks represents a distinct assessment item or item information. The first line of each codebook entry contains the following information:

SEQ. NO.	Sequence number. In conjunction with the numbers assigned in the layouts, the fields in the codebooks are numbered sequentially.
FIELD NAME	A brief description of the information in the field. The field names in the codebooks correspond to those in the layouts.
REL. IND.	Released indicator. Indicates that an item has been released to the public (R) or that an item has been held secure (U).
TYPE	In conjunction with the three field types defined for the layouts above, the field type is designated as Continuous (C), Discrete (D), or Discrete with "I Don't Know" (DI).
BLOCK	For assessment items, indicates the block in which an item appeared for the grade/age group of students for which the codebook was prepared. (See Table 8-2 for block designations.)
ITEM NO.	Indicates the number of an item within a block for the grade/age group of students for which the codebook was prepared.

Table 8-2

Block Designations and File Locations

BLOCK	LOCATION	DESCRIPTION
ID	Student Files	Assessment identification and information from booklet cover
WT	Student Files	Weights and derived variables
BA	Student Files	Common background items
CA-CF*	Student Files	Computer competence items
HA-HD*	Student Files	U.S. history items
LA-LD*	Student Files	Literature items
MA-MK*	Student Files	Mathematics items
RA-RF*	Student Files	Reading items
SA-SK*	Student Files	Science items
XI	Excluded Student Files	Assessment identification and information from questionnaire cover
XW	Excluded Student Files	Weights and derived variables
XQ	Excluded Student Files	Excluded student questionnaire items
TI	Teacher Files	Assessment identification and information from questionnaire cover
TW	Teacher Files	Weights and derived variables
TQ	Teacher Files	Teacher questionnaire items
SS	School Files	Assessment identification
SW	School Files	Weights and derived variables
SQ	School Files	School questionnaire items
TC	School Files	Computer coordinator questionnaire items

*To save file space, the numeric block designation (1 through 11) has been replaced by an alphabetic designation (A through K).

AGES	Indicates the student grade/age or age groups to whom an item was administered, as follows:

Spiral Samples	Bridge Samples
1 = Grade 3/Age 9	1 = Age 9
2 = Grade 7/Age 13	2 = Age 13
3 = Grade 11/Age 17	3 = Age 17

NAME/DESCRIPTION	Provides a brief description of the item or information in the field.
OLD ID (NAEP ID)	If an item has been used in the past, the ID number assigned to it during previous assessments is included as the last number on the first line of the codebook entry.

Please note that, for items administered to more than one grade/age group, the information under BLOCK and ITEM NO. refers only to the grade/age group for which a codebook was prepared. For example, the block location and item number for an item in the grade 3/age 9 (grade/age 1) codebook apply to grade/age 1 only. The block location and item number may be different for grade/age groups 2 and 3. The block and item numbers for each grade/age group are indicated in the appropriate codebooks.

For all discrete variables and items, the third and subsequent lines contain each valid data value, its associated label, and the unweighted frequency of that value in the data file associated with the codebook. (For achievement items, the key data value is indicated by an asterisk.) The last line under each discrete variable entry contains the "TOTAL" or sum of the frequency counts as an extra check for analyses.

8.5 MACHINE-READABLE CATALOGS

Each raw data file has a corresponding machine-readable file; these files all have the suffix CAT. The files, designed to help users extract data of interest from the raw data, contain a record for each variable or item on the raw data files. Table 8-3 contains the machine-readable catalog data layout. Specific information concerning the contents of the catalogs is provided below.

FIELD SEQUENCE NUMBER	Fields are numbered sequentially to represent the order in which they appear on the raw data record.
FIELD NAME	A short name of up to eight characters that identifies the field.
START COLUMN	The start location of the field on the raw data record.
END COLUMN	The end location of the field on the raw data record.
FIELD WIDTH	The number of characters in the field.

Table 8-3
Machine-Readable Catalog Data Layout

START END COLUMNS	FIELD WIDTH	FIELD DESCRIPTION	COMMENTS
1 - 4	4	Field Sequence Number	
5 - 12	8	Field Name	New NAEP Identification
13 - 16	4	Start Column	
17 - 20	4	End Column	
21 - 22	2	Field Width	
23 - 23	1	Decimal Places	
24 - 24	1	Field Type	1=Discrete 2=Continuous
25 - 27	3	Minimum Valid Response	Excluding I Don't Know responses
28 - 30	3	Maximum Valid Response	
31 - 32	2	Correct Response	
33 - 34	2	I Don't Know (IDK) Code	
35 - 36	2	No Response Code	
37 - 38	2	Multiple and Out-Of-Range Code	
39 - 88	50	Field Label	
89 - 104	16	Old NAEP Identification	
105 - 128	24	IRT Parameters	Three fields, eight positions each
129 - 130	2	Number of Response Values and Labels	This number determines the number of response values and labels that follow
131 - 132	2	Response Value	1st Value
133 - 152	20	Response Label	
153 - 154	2	Response Value	2nd Value
155 - 174	20	Response Label	
. . .		. . .	. . .
549 - 550	2	Response Value	20th Value
551 - 570	20	Response Label	

DECIMAL PLACES	The number of digits to the right of the decimal point in the field. The raw data contain implicit decimal points.
FIELD TYPE	The files include two field types:
Type 1	(Discrete) fields designate discrete data with a fixed number of responses. Type 1 fields may include raw item responses or imputed categorical variables.
Type 2	(Continuous) fields designate continuous numerical data without fixed ranges.
MINIMUM VALID RESPONSE	The minimum value of valid responses for an item, excluding "I Don't Know" response.
MAXIMUM VALID RESPONSE	The maximum value of valid responses for an item, excluding "I Don't Know" response.
CORRECT RESPONSE	This is a two-position field. If the second position is blank, the item has a unique correct response in the first position. If the second position is not blank, the item has more than one correct response and the field indicates the correct response range. The first position is the lowest correct response value; the second is the highest correct response value. For example, if possible responses for a professionally scored reading item ranged from 0 to 5, and 3 to 5 was considered an acceptable response, the first position of the field would contain a 3 and the second position would contain a 5.
I DON'T KNOW	For appropriate multiple-choice items, the numbers in the "IDK Value" column reflect the option to respond by indicating "I Don't Know." "I Don't Know" values have been recoded to 7 or 77. Otherwise, this field is blank.
NO RESPONSE CODE	Where applicable, omitted or not reached values have been recoded to 8 or 88. Otherwise, this field is blank.
MULTIPLE AND OUT-OF-RANGE CODE	Where applicable, multiple or out-of-range responses have been recoded to 9 or 99. Otherwise, this field is blank.
OLD NAEP ID	If an item has been used in the past, this is the number assigned to it during previous assessments.

IRT PARAMETERS	Three eight-character fields containing IRT item parameters: "a" (discrimination); "b" (difficulty); and "c" (lower asymptote). Each parameter is represented to a precision of five decimal places with an explicit decimal point. (Not included on Version 1 tapes.)
NO. OF RESPONSE VALUES & LABELS	The number of possible valid responses plus "I Don't Know," no response, multiple and out-of-range.
RESPONSE VALUES AND LABELS	For each possible item response, a two-position field which indicates the response value and a 20-position text field which provides a brief description of the response.

8.6 SPSS-X AND SAS CONTROL STATEMENT FILES

All data files on the public-use data tapes are accompanied by separate control files to facilitate the creation of SPSS-X and SAS system files. These control files include statements for variable definitions, variable labels, missing value codes, value labels, and an optional section for creating and storing scored variables. Each set of control statements also generates unweighted descriptive statistics of the reporting variables for the related data file and a listing of the contents of the saved system file.

Specific details on the structure and use of these control files are provided in Chapter 9.

Table 8-4
1985-86 Public-Use Data Tape Description

Grade 3/Age 9

DATA FILES	RECORD LENGTH	BLOCK SIZE	# OF RECORDS	FILE NAME
Data Files				
1. Student Spiral Data	1098	18666	21287	Y17RMSC1.BIB.DAT
2. Bridge Booklet 1	544	19040	2315	Y17RMSC1.BR1.DAT
3. Bridge Booklet 2	537	18795	2361	Y17RMSC1.BR2.DAT
4. Bridge Booklet 3	550	18700	2256	Y17RMSC1.BR3.DAT
5. Bridge Booklet 4	557	18938	1994	Y17RMSC1.BR4.DAT
6. Bridge Booklet 5	540	18900	2048	Y17RMSC1.BR5.DAT
7. School Questionnaire	549	18666	632	Y17RMSC1.SCQ.DAT
SPSS-X Control Statement Files				
8. Student Spiral Data	80	19040	3890	Y17RMSC1.BIB.SPX
9. Bridge Booklet 1	80	19040	865	Y17RMSC1.BR1.SPX
10. Bridge Booklet 2	80	19040	842	Y17RMSC1.BR2.SPX
11. Bridge Booklet 3	80	19040	902	Y17RMSC1.BR3.SPX
12. Bridge Booklet 4	80	19040	991	Y17RMSC1.BR4.SPX
13. Bridge Booklet 5	80	19040	866	Y17RMSC1.BR5.SPX
14. School Questionnaire	80	19040	723	Y17RMSC1.SCQ.SPX
SAS Control Statement Files				
15. Student Spiral Data	80	19040	4021	Y17RMSC1.BIB.SAS
16. Bridge Booklet 1	80	19040	946	Y17RMSC1.BR1.SAS
17. Bridge Booklet 2	80	19040	897	Y17RMSC1.BR2.SAS
18. Bridge Booklet 3	80	19040	966	Y17RMSC1.BR3.SAS
19. Bridge Booklet 4	80	19040	1027	Y17RMSC1.BR4.SAS
20. Bridge Booklet 5	80	19040	912	Y17RMSC1.BR5.SAS
21. School Questionnaire	80	19040	1260	Y17RMSC1.SCQ.SAS
Machine-Readable Catalog Files				
22. Student Spiral Data	570	18810	765	Y17RMSC1.BIB.CAT
23. Bridge Booklet 1	570	18810	213	Y17RMSC1.BR1.CAT
24. Bridge Booklet 2	570	18810	207	Y17RMSC1.BR2.CAT
25. Bridge Booklet 3	570	18810	218	Y17RMSC1.BR3.CAT
26. Bridge Booklet 4	570	18810	226	Y17RMSC1.BR4.CAT
27. Bridge Booklet 5	570	18810	210	Y17RMSC1.BR5.CAT
28. School Questionnaire	570	18810	382	Y17RMSC1.SCQ.CAT

Table 8-5
1985-86 Public-Use Data Tape Description

Grade 7/Age 13

DATA FILES	RECORD LENGTH	BLOCK SIZE	# OF RECORDS	FILE NAME
Data Files				
1. Student Spiral Data	1606	17666	27668	Y17RMSC2.BIB.DAT
2. Bridge Booklet 1	597	18507	2075	Y17RMSC2.BR1.DAT
3. Bridge Booklet 2	567	18711	2054	Y17RMSC2.BR2.DAT
4. Bridge Booklet 3	585	18720	2071	Y17RMSC2.BR3.DAT
5. Bridge Booklet 4	608	18848	2032	Y17RMSC2.BR4.DAT
6. Bridge Booklet 5	584	18688	2146	Y17RMSC2.BR5.DAT
7. School Questionnaire	553	18802	567	Y17RMSC2.SCQ.DAT
SPSS-X Control Statement Files				
8. Student Spiral Data	80	19040	6804	Y17RMSC2.BIB.SPX
9. Bridge Booklet 1	80	19040	1131	Y17RMSC2.BR1.SPX
10. Bridge Booklet 2	80	19040	1005	Y17RMSC2.BR2.SPX
11. Bridge Booklet 3	80	19040	1115	Y17RMSC2.BR3.SPX
12. Bridge Booklet 4	80	19040	1299	Y17RMSC2.BR4.SPX
13. Bridge Booklet 5	80	19040	1120	Y17RMSC2.BR5.SPX
14. School Questionnaire	80	19040	731	Y17RMSC2.SCQ.SPX
SAS Control Statement Files				
15. Student Spiral Data	80	19040	7041	Y17RMSC2.BIB.SAS
16. Bridge Booklet 1	80	19040	1247	Y17RMSC2.BR1.SAS
17. Bridge Booklet 2	80	19040	1093	Y17RMSC2.BR2.SAS
18. Bridge Booklet 3	80	19040	1222	Y17RMSC2.BR3.SAS
19. Bridge Booklet 4	80	19040	1359	Y17RMSC2.BR4.SAS
20. Bridge Booklet 5	80	19040	1194	Y17RMSC2.BR5.SAS
21. School Questionnaire	80	19040	1273	Y17RMSC2.SCQ.SAS
Machine-Readable Catalog Files				
22. Student Spiral Data	570	18810	1270	Y17RMSC2.BIB.CAT
23. Bridge Booklet 1	570	18810	264	Y17RMSC2.BR1.CAT
24. Bridge Booklet 2	570	18810	236	Y17RMSC2.BR2.CAT
25. Bridge Booklet 3	570	18810	253	Y17RMSC2.BR3.CAT
26. Bridge Booklet 4	570	18810	276	Y17RMSC2.BR4.CAT
27. Bridge Booklet 5	570	18810	253	Y17RMSC2.BR5.CAT
28. School Questionnaire	570	18810	386	Y17RMSC2.SCQ.CAT

Table 8-6
1985-86 Public-Use Data Tape Description

Grade 11/Age 17

DATA FILES	RECORD LENGTH	BLOCK SIZE	# OF RECORDS	FILE NAME
Data Files				
1. Student Spiral Data	2407	16849	39753	Y17RMSC3.BIB.DAT
2. Bridge Booklet 4	631	18930	1934	Y17RMSC3.BR4.DAT
3. Bridge Booklet 5	615	19065	1934	Y17RMSC3.BR5.DAT
4. School Questionnaire	598	18538	433	Y17RMSC3.SCQ.DAT
SPSS-X Control Statement Files				
5. Student Spiral Data	80	19040	10489	Y17RMSC3.BIB.SPX
6. Bridge Booklet 4	80	19040	1320	Y17RMSC3.BR4.SPX
7. Bridge Booklet 5	80	19040	1209	Y17RMSC3.BR5.SPX
8. School Questionnaire	80	19040	767	Y17RMSC3.SCQ.SPX
SAS Control Statement Files				
9. Student Spiral Data	80	19040	11196	Y17RMSC3.BIB.SAS
10. Bridge Booklet 4	80	19040	1421	Y17RMSC3.BR4.SAS
11. Bridge Booklet 5	80	19040	1317	Y17RMSC3.BR5.SAS
12. School Questionnaire	80	19040	1329	Y17RMSC3.SCQ.SAS
Machine-Readable Catalog Files				
13. Student Spiral Data	570	18810	2069	Y17RMSC3.BIB.CAT
14. Bridge Booklet 4	570	18810	297	Y17RMSC3.BR4.CAT
15. Bridge Booklet 5	570	18810	281	Y17RMSC3.BR5.CAT
16. School Questionnaire	570	18810	410	Y17RMSC3.SCQ.CAT

Chapter 9

WORKING WITH SPSS-X AND SAS

Chapter 9: WORKING WITH SPSS-X AND SAS[1]

9.1 INTRODUCTION

This chapter discusses the use of the statistical software SPSS-X and SAS in analyzing 1985-86 NAEP data. Included are procedures for creating SPSS-X and SAS system files, merging files using SPSS-X and SAS, and using the jackknife procedure with SPSS-X and SAS to estimate standard errors.

9.2 SPSS-X AND SAS CONTROL STATEMENT FILES

All data files on the public-use data tapes are accompanied by separate control files to facilitate the creation of SPSS-X and SAS system files. These control files include statements for variable definitions, variable labels, missing value codes, value labels, and an optional section for creating and storing scored variables. Each set of control statements also generates unweighted descriptive statistics of the reporting variables for the related data file and a listing of the contents of the saved system file.

The common features of both types of control files, as well as general guidelines, are provided below.

VARIABLE DEFINITION	The field names are listed in the order in which they appear on the file, along with their column position and input formats. If the field is numeric with no decimal places, no format is provided. Otherwise the format is indicated by a number for the number of decimal places, or a '$' or '(A)' for a non-numeric field.
VARIABLE LABELS	A 40-character text description for each field.
MISSING VALUES	All blank fields in the data are automatically set to the system missing value by each package. However, all multiple-choice and some open-ended items were prone to either multiple or out-of-range responses. These items were coded as fields of 9s in the data files. The control statement files instruct each system to treat these values as missing.

---

[1] Data collected from the Teacher Questionnaire and the Excluded Student Questionnaire, sampling weights for schools, excluded students, and teachers, and IRT parameters for student responses are not contained on the Version 1.0 data tapes. They will be included in Version 2.0.

VALUE LABELS	All numeric fields with discrete (or categorical) values are provided with 20-character text descriptors for each value within the variable's range.
SCORING	For each item with one or more correct responses, control statements are provided for creating a scored variable, its label, and its value labels. The scoring of each item is performed according to the following scheme: missing values are copied as is, correct response values are recoded to 1, all other values, including No Response and "I Don't Know," are recoded as 0. The scoring of the No Response and "I Don't Know" values are coded separately from other incorrect responses to allow the user to edit these control statements and substitute alternate values. The scored variable names are derived by replacing the value in the sixth and seventh digits of the original name with that value plus 50. For example, the scored version of item number N001503A is N001553A. The entire block of scoring control statements is performed conditionally and, by default, will not be saved on the system file. The user must edit only one statement in either type of file to invoke the scoring and to save the new variables.

IMPORTANT NOTE: The system file generation programs cannot run if both the control statement file and its corresponding data file reside on the same tape. Both SPSS-X and SAS will try to read a data file before they have completed processing the control statement file, which is physically impossible if both files are on the same tape. The user is advised to copy the control files to disk, as they require less storage space and allow the user to edit the control statements before generating the system files.

9.3 CREATING SPSS-X SYSTEM FILES

Each SPSS-X control statement file is linked to its corresponding data file through the file name: the suffix DAT in the data file name is replaced by SPX to obtain the control statement file name. For example, file "Y17RDWR1.BIB.SPX" is the control statement file for data file "Y17RDWR1.BIB.DAT."

All SPSS-X control statement files have been generated according to the structure in Table 9-1.

The TEMPORARY command instructs SPSS-X to perform the subsequent scoring statements on a temporary basis and delete the new variables after the next procedure encountered (FREQUENCIES). Thus, the scored variables will NOT be saved on the system file unless the TEMPORARY command is commented or edited out.

Table 9-1
SPSS-X Control Statement Synopsis

```
TITLE
   label for sysout of file generation run
FILE LABEL
   label to be stored with file
DOCUMENT
   text description of data to be saved in file
DATA LIST FILE=RAWDATA/
   variable names, locations, & formats
VARIABLE LABELS
   40-character label for each variable
MISSING VALUES
   list of variables to have user-missing values assigned
VALUE LABELS
   variable names, values, & value labels
TEMPORARY                          [delete this statement to
                                    save scored variables]
RECODE
   oldvar (SYSMIS=SYSMIS)  (9=9)  (keyval=1)
   (nrval=0)  (idkval=0)  (ELSE=0)  INTO  newvar
 .
 .
 .

VARIABLE LABELS
   labels for new variables
MISSING VALUES
   for multiple & out-of-range recodes
 VALUE LABELS
   1=Correct    0=Incorrect
 FREQUENCIES
   reporting variables
 SAVE   OUTFILE=SYSFILE/COMPRESSED
 DISPLAY LABELS
```

All control statement files assume that the file handle (or DDNAME) for the input data file is RAWDATA, and the file handle for the output system file is SYSFILE.

The control statements were coded according to the command and procedure descriptions in the SPSS-X User's Guide (SPSS Inc., McGraw-Hill Book Company, 1983). They were tested under SPSS-X Version 2.1 (IBM-OS/MVS).

9.4 CREATING SAS SYSTEM FILES

Each SAS control statement file is linked to its corresponding data file through the file name: the suffix DAT in the data file name is replaced by SAS to obtain the control statement file name. For example, file "Y17RDWR1.BIB.SAS" is the control statement file for data file "Y17RDWR1.BIB.DAT."

All SAS control statement files have been generated according to the structure in Table 9-2.

All SAS control files use the SAS Macro Language facility to reduce the number of source statements generated and to provide consistent performance of repetitive functions. Therefore, the user must ensure that the MACRO option is invoked whenever processing any of the control statement files.

The DO OVER through END statements following each ARRAY statement set up the conversion of the "I Don't Know," No Response, and Multiple Response codes to the system missing value. However, once this conversion is executed and saved on the system file, these recoded values will be indistinguishable from actual missing values on the original data file. For this reason, these statements have been commented out to allow the user to decide which, if any, of the values are to be recoded. To activate the recoding, delete the asterisks preceding the DO OVER and END statements, and from the appropriate IF THEN statement(s).

The SCORE MACRO sequence sets up the creation of scored variables and their labels. To save these variables on the system file, edit the statement immediately following the MACRO and delete the asterisk in the first column.

The control statements were coded according to the command and procedure descriptions in the SAS User's Guide: Basics 1982 Edition (SAS Institute Inc., 1982). They were tested under SAS Version 82.4 (IBM-OS/MVS).

Table 9-2

SAS Control Statement Synopsis

```
TITLE
DATA SYSFILE.xxx;
INFILE RAWDATA;
INPUT
  Variable names, column positions, & formats
LABEL
  40-character variable labels
ARRAY MVn (I)
  list of variables to be recoded for missing
*DO OVER MVn;
  *  IF MVn=7   THEN MVn=.;
  *  IF MVn=8   THEN MVn=.;
  *  IF MVn=9   THEN MVn=.;
  *  END;
%MACRO FMTVAR;
 PROC FORMAT;
 VALUE
  variable name    value=value label
  .
  .
%MEND FMTVAR;
LENGTH DEFAULT=2
  other variables with appropriate lengths;
%MACRO SCORE;
%MACRO RECODE;
  SAS macro to perform scoring for each variable
%MEND RECODE;
%RECODE
  (oldvar,newvar,idkval,nrval,keyval)
  .
  .
LABEL
  labels for new variables
PROC FORMAT;
  value labels for new variables
%MEND SCORE;
*%SCORE;                                   [delete asterisk to
                                            save scored variables]
%FMTVAR;
PROC FREQ;
TABLES
  reporting variables
PROC CONTENTS NOSOURCE POSITION;
```

9.5 MERGING FILES UNDER SPSS-X OR SAS

The NAEP data files are structured to facilitate matching among the four instruments (Student, Teacher, School, and Excluded Student). However, for the purposes of analysis and reporting, only four types of linkages are valid:

1) School with Student (spiral or bridge)
2) Teacher with Student (spiral only)
3) School with Excluded Student (spiral or bridge)
4) School with Teacher with Student (spiral only)

The primary linkage on all files is through the scrambled PSU and school code fields: SCRPSU, SSCRPSU, TSCRPSU and XSCRPSU. The secondary linkage for the BIB Spiral and Teacher files is through the scrambled teacher code fields: SCRTC and TSCRTC. All files are sorted by these fields to permit direct match-merging without the need to re-sort.

When a hierarchical file match is performed, both SPSS-X and SAS build a rectangular file at the level of the lowest file in the match. Each record from the higher order file is repeated across the corresponding records of the lower order file. For example, in matching School with Student data, the information from one school record is repeated across all student records belonging to that school. Clearly, the number of variables from the higher order file will have a greater impact on the size of the resulting merged file.

The examples contained in Tables 9-3 through 9-6 will perform direct matches according to the four linkages listed above. The KEEP statements are not necessary to the performance of the merge, but when applied to only those variables required for analysis will make more efficient use of computer resources. These examples also assume that no transformations are to be performed on the input files. If transformations are desired for analysis, the most efficient course to follow would be to transform the variables from the higher order file first, perform the match procedure, then transform the variables from the lower order file.

9.6 COMPUTING THE ESTIMATED VARIANCE OF A MEAN (JACKKNIFING) USING SPSS-X OR SAS

A method for computing the estimated variance of a mean is presented in this section in SPSS-X and SAS program code form (see Section 7.3.1 in Chapter 7 for a discussion of the jackknife procedure). This method, the multi-weight method, may be used for any variable.

For each variable to be jackknifed, generate two vectors of weighted sums and products. Sum these vectors across the entire file using the AGGREGATE (SPSS-X) or SUMMARY (SAS) procedures. From the weighted sums compute the weighted means and thence compute the estimated variance and standard error.

One advantage to this approach is that it will accomplish the computation in one pass of the data. Another advantage, afforded by the AGGREGATE (SPSS-X) and SUMMARY (SAS) procedures, is the facility to compute subgroup statistics by using the BREAK keyword (SPSS-X) or CLASS option (SAS) with the variable(s) defining the subgroups. All computations performed subsequent to the aggregation procedure are performed on each record of the collapsed file, corresponding to one of the subgroups. In the examples in Tables 9-7 and 9-8, the variable SEX is used as a break control variable, and the derived statistics are printed for each sex code.

Table 9-3

Matching School and Student Files

```
                              SPSS-X

MATCH FILES
  TABLE=SCHOOL/
     RENAME=(SSCRPSU=SCRPSU)/
  FILE=STUDENT/
  KEEP=SCRPSU,other school & student variables/
  BY=SCRPSU
```

```
                              SAS

DATA MATCH1;
  MERGE SCHOOL (RENAME=(SSCRPSU=SCRPSU)
                KEEP=SSCRPSU other school variables)
        STUDENT(KEEP=SCRPSU other student variables);
  BY SCRPSU;
```

Table 9-4

Matching Teacher and Student Files

```
                              SPSS-X

MATCH FILES
  TABLE=TEACHER/IN=INTEACH/
     RENAME=(TSCRPSU=SCRPSU)(TSCRTC=SCRTC)/
  FILE=STUDENT/
  KEEP=SCRPSU,SCRTC,other teacher & student variables/
  BY=SCRPSU,SCRTC
SELECT IF (INTEACH)
```

```
                              SAS

DATA MATCH2;
  MERGE TEACHER(RENAME=(TSCRPSU=SCRPSU TSCRTC=SCRTC) IN=T
                KEEP=TSCRPSU TSCRTC other teacher variables)
        STUDENT(KEEP=SCRPSU SCRTC other student variables);
  BY SCRPSU SCRTC;
IF T;
```

Table 9-5

Matching School and Excluded Student Files

SPSS-X

```
MATCH FILES
  TABLE=SCHOOL/
     RENAME=(SSCRPSU=SCRPSU)/
  FILE=EXCLUDE/
     RENAME=(XSCRPSU=SCRPSU)/
  KEEP=SCRPSU,other school & excluded student variables/
  BY=SCRPSU
```

SAS

```
DATA MATCH3;
  MERGE SCHOOL (RENAME=(SSCRPSU=SCRPSU)
           KEEP=SSCRPSU other school variables)
        EXCLUDE(RENAME=(XSCRPSU=SCRPSU)
           KEEP=SCRPSU other excluded student variables);
  BY SCRPSU;
```

Table 9-6

Matching School, Teacher and Student Files

SPSS-X

```
MATCH FILES
  TABLE=SCHOOL/IN=INSCHOOL/
     RENAME=(SSCRPSU=SCRPSU)/
  FILE=TEACHER/
     RENAME=(TSCRPSU=SCRPSU)(TSCRTC=SCRTC)/
  KEEP=SCRPSU,SCRTC,other school & teacher variables/
  BY=SCRPSU
SELECT IF (INSCHOOL)
MATCH FILES
  TABLE=*/IN=INTEACH/
  FILE=STUDENT/
  KEEP=SCRPSU,SCRTC,other school,teacher & student variables/
  BY=SCRPSU,SCRTC
SELECT IF (INTEACH)
```

SAS

```
DATA MATCH4A;
  MERGE SCHOOL (RENAME=(SSCRPSU=SCRPSU) IN=S
              KEEP=SSCRPSU other school variables)
        TEACHER(RENAME=(TSCRPSU=SCRPSU TSCRTC=SCRTC)
              KEEP=TSCRPSU TSCRTC other teacher variables);
  BY SCRPSU;
  IF S;
DATA MATCH4B;
  MERGE MATCH4A (IN=T
               KEEP=SCRPSU SCRTC other school
                                      & teacher variables)
        STUDENT(KEEP=SCRPSU SCRTC other student variables);
  BY SCRPSU SCRTC;
  IF T;
```

Table 9-7

Standard Error Computation: Multi-Weight Method Using SPSS-X

SPSS-X

```
GET FILE=STUDENT/
   KEEP=SEX,WEIGHT,JKWT01 TO JKWT38,X
VECTOR WT=JKWT01 TO JKWT38
VECTOR WX(38)
SELECT IF (NOT SYSMIS(X))
COMPUTE WTX=WEIGHT*X
LOOP #I=1 TO 38
.  COMPUTE WX(#I) = WT(#I)*X
END LOOP
AGGREGATE  OUTFILE=*/BREAK=SEX/UWN=N(WEIGHT)/
   SWT,SW1 TO SW38 = SUM(WEIGHT,JKWT01 TO JKWT38)/
   SWX,SX1 TO SX38 = SUM(WTX,WX1 TO WX38)/
VECTOR SW = SW1 TO SW38
VECTOR SX = SX1 TO SX38
COMPUTE XBAR = SWX/SWT
COMPUTE XVAR = 0
LOOP #I=1 TO 38
.  COMPUTE #DIFF = SX(#I)/SW(#I) - XBAR
.  COMPUTE XVAR = XVAR + #DIFF * #DIFF
END LOOP
COMPUTE XSE = SQRT(XVAR)
PRINT FORMATS XVAR,XSE (F8.4)
LIST VARIABLES=SEX,UWN,SWT,XBAR,XVAR,XSE
FINISH
```

Table 9-8

Standard Error Computation: Multi-Weight Method Using SAS

SAS

```
DATA A;
   SET  STUDENT;
   ARRAY WT JKWT01-JKWT38;
   ARRAY WX WX1-WX38;
   IF (X NE .);
   WTX = WEIGHT*X;
   DO OVER WT;
      WX = WT*X;
   END;
PROC SUMMARY;
   CLASS SEX;
   VAR WEIGHT JKWT01-JKWT38 WTX WX1-WX38;
   OUTPUT OUT=B    N(WEIGHT)=UWN
      SUM(WEIGHT WTX JKWT01-JKWT38 WX1-WX38)=
          SWT SWX SW1-SW38 SX1-SX38;
DATA C;
   SET B;
   ARRAY SW SW1-SW38;
   ARRAY SX SX1-SX38;
   XBAR = SWX/SWT;
   XVAR = 0;
   DO OVER SW;
      DIFF = (SX/SW)-XBAR;
      XVAR = XVAR+DIFF*DIFF;
   END;
   XSE = SQRT(XVAR);
PROC PRINT;
   VAR SEX UWN SWT XBAR XVAR XSE;
```

Appendix A

NAEP HISTORY

Appendix A: NAEP HISTORY

The National Assessment of Educational Progress (NAEP) is a continuing, congressionally mandated national survey of the knowledge, skills, understandings, and attitudes of young Americans in major learning areas usually taught in school. Its primary goals are to detect and report the current status of, as well as changes in, the educational attainments of young Americans, and to report long-term trends in those attainments. The purpose of NAEP is to gather information which will aid educators, legislators and others in improving the educational experience of youth in the United States. It is the first ongoing effort to obtain comprehensive and dependable achievement data on a national basis in a uniform, scientific manner.

Between 1964 and 1969, initial assessment planning and development activities were conducted for NAEP with support from both the Carnegie Corporation and the Ford Foundation. During this time, objectives and exercises were developed for many of the learning areas, sampling and data collection strategies were planned, and data analysis plans were formulated and outlined.

From its inception, NAEP has developed assessments through a consensus process. In the process, educators, scholars, and lay persons design objectives for each learning area, proposing general goals they think Americans should achieve in the course of their education. After careful reviews, the objectives are given to item writers, who develop measurement instruments appropriate to the objectives.

After the items pass extensive reviews by subject-matter specialists, measurement experts, and lay persons, and are tested in schools throughout the country, they are administered to a stratified multistage national probability sample. The young people sampled are selected so that assessment results may be generalized to the entire national population.

NAEP collected data for the first time in 1969. Since that time, samples have included over one million 9-, 13- and 17-year-old students and, as funding would allow, 17-year-olds who had left school and adults 26 to 35 years of age. In 1983-84, grade samples of students were added to the assessment. As Table A-1 illustrates, assessments have focused on traditional learning areas such as reading, writing, mathematics, and science, and on less traditional areas such as citizenship, art, literature, music, and career and occupational development.

Since 1971, NAEP has been solely supported by federal funds. Funding agencies have included the Office of Education, the National Center for Education, and the National Institute of Education. NAEP is currently supported by the Office for Educational Research and Improvement, Center for Statistics.

NAEP was administered by the Education Commission of the States (ECS) through 1982. In 1983, Educational Testing Service (ETS) assumed responsibility for administration of the project. In assuming responsibility for NAEP, ETS has incorporated an updated sampling design and, at the same time, has made a concerted effort to ensure continuity with previous assessments.

Public-use data tapes were first produced in 1975, allowing outside researchers access to the NAEP database. In June 1985 ETS produced its first public-use data tapes, in a new format, for the 1983-84 assessment. Although it will not be possible to change the format of the tapes issued for assessments prior to 1983-84, the new format produced by ETS makes the tapes easier to use (e.g., files have been more simply organized, documentation has been improved and made more accessible).

Table A-1
National Assessment of Educational Progress
Learning Areas, Grades, and Ages Assessed
1969-1986

ASSESSMENT YEAR	LEARNING AREAS	GRADES/AGES ASSESSED*									
		Grade 3	Grade 4	Age 9	Grade 7	Grade 8	Age 13	Grade 11	Age 17IS	Age 17OS	Age ADULT
Year 1/1969-70	Science			X			X		X	X	X
	Writing			X			X		X	X	X
	Citizenship			X			X		X	X	X
Year 2/1970-71	Reading			X			X		X	X	X
	Literature			X			X		X	X	X
Year 3/1971-72	Music			X			X		X	X	X
	Social Studies			X			X		X	X	X
Year 4/1972-73	Science (2)			X			X		X	X	X
	Mathematics			X			X		X	X	X
Year 5/1973-74	Career and Occupational Development			X			X		X	X	X
	Writing (2)			X			X		X	X	
Year 6/1974-75	Reading (2)			X			X		X	X	
	Art			X			X		X	X	
Year 7/1975-76	Citizenship/ Social Studies (2)			X			X		X	X	
	Mathematics**								X	X	
Year 8/1976-77	Science (3)			X			X		X		
	Basic Life Skills**								X		
	Health**									X	
	Energy**									X	
	Reading** (3)									X	
	Science** (3)									X	
Year 9/1977-78	Mathematics (2)			X			X		X		
	Consumer Skills**								X		
Year 10/1978-79	Art (2)			X			X		X		
	Music (2)			X			X		X		
	Writing (3)			X			X		X		
Year 11/1979-80	Reading (4)			X			X		X	X	
	Literature (2)			X			X		X	X	

Table A-1
(continued)

ASSESSMENT YEAR	LEARNING AREAS	GRADES/AGES ASSESSED*									
		Grade 3	Grade 4	Age 9	Grade 7	Grade 8	Age 13	Grade 11	Age 17IS	Age 17OS	Age ADULT
Year 12/1980-81	No Data Collection										
Year 13/1981-82	Mathematics (3)			X			X		X		
	Citizenship/ Social Studies (3)			X			X		X		
	Science** (4)			X			X		X		
Year 14/1982-83	No Data Collection										
Year 15/1983-84	Reading (5)		X	X		X	X	X	X		
	Writing (4)		X	X		X	X	X	X		
Year 16/1984-85	Literacy**										X
Year 17/1985-86	Mathematics (4)	X		X	X		X	X	X		
	Reading (6)	X		X	X		X	X	X		
	Science (5)	X		X	X		X	X	X		
	Computer Competence	X		X	X		X	X	X		
	Literature**							X	X		
	U. S. History**							X	X		

* 17IS denotes 17-year-olds enrolled in public or private schools; 17OS denotes 17-year-olds who dropped out of school or graduated prior to the assessment

** Small, special-interest assessment conducted on limited samples at specific ages

( ) Indicates second and subsequent assessments of a learning area

Appendix B

RELATED MACHINE-READABLE DATA FILES AND PRINTED REPORTS

APPENDIX B

Related Machine-Readable Data Files

Data files are also available for the following NAEP assessments:

YEAR	GRADE/AGE	LEARNING AREA(S)
1985	Adult	Literacy
1983-84	Grade 4/Age 9	Reading and Writing
1983-84	Grade 8/Age 13	Reading and Writing
1983-84	Grade 11/Age 17	Reading and Writing
1981-82	Age 9	Math and Citizenship/Social Studies
1981-82	Age 13	Math and Citizenship/Social Studies
1981-82	Age 17	Math and Citizenship/Social Studies
1981-82	Age 9, 13, 17	Science Probe
1979-80	Age 9	Reading/Literature
1979-80	Age 13	Reading/Literature
1979-80	Age 17	Reading/Literature
1978-79	Age 9	Art, Music, and Writing
1978-79	Age 13	Art, Music, and Writing
1978-79	Age 17	Art, Music, and Writing
1978-79	Age 9, 13, 17	Writing Essay Responses
1978	Age 17	Consumer Skills
1977-78	Age 9	Mathematics
1977-78	Age 13	Mathematics
1977-78	Age 17	Mathematics
1977	Age 17	Basic Life Skills
1977	Adults	Science, Energy, Reading, and Health
1976-77	Age 9	Science
1976-77	Age 13	Science
1976-77	Age 17	Science
1975-76	Age 9	Citizenship/Social Studies
1975-76	Age 13	Citizenship/Social Studies
1975-76	Age 17	Citizenship/Social Studies
1975-76	Age 13, 17	Supplemental Mathematics
1974-75	Age 9	Reading
1974-75	Age 13	Reading
1974-75	Age 17	Reading
1974-75	Age 9	Art
1974-75	Age 13	Art
1974-75	Age 17	Art

(continued)

YEAR	GRADE/AGE	LEARNING AREA(S)
1973-74	Age 9	Career & Occupational Development and Writing
1973-74	Age 13	Career & Occupational Development and Writing
1973-74	Age 17	Career & Occupational Development and Writing
1973-74	Adults	Career & Occupational Development and Writing
1973-74	Age 9, 13, 17	Writing Essay Responses
1972-73	Age 9	Science and Mathematics
1972-73	Age 13	Science and Mathematics
1972-73	Age 17	Science and Mathematics
1972-73	Adult	Science and Mathematics
1970-71	Age 9	Reading/Literature
1970-71	Age 13	Reading/Literature
1970-71	Age 17	Reading/Literature
1970-71	Adults	Reading/Literature
1969-70	Age 9, 13, 17	Writing Essay Responses

For more information about NAEP public-use data tapes, contact:

Norma Norris
Senior Research Data Analyst
National Assessment of Educational Progress
Educational Testing Service, 22-T
Princeton, NJ 08541
(609) 734-5898

Related Printed Reports

The following reports provide additional information about the National Assessment of Educational Progress. Please note that direct sources are included for each publication, but that most may be obtained through the Educational Resources Information Center (ERIC) database.

MATHEMATICS

Education Commission of the States. (1981). *Mathematics objectives, 1981-82 assessment.* (ED 211 352) (NAEP-13-MA-10) (ISBN-0-89398-147-8) Denver, CO: National Assessment of Educational Progress, 41p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1979). *The second assessment of mathematics, 1977-78, released exercise set.* (ED 187 543) Denver, CO: National Assessment of Educational Progress, 366p.

Education Commission of the States. (1978). *Mathematics objectives: Second assessment.* (ED 156 439) Denver, CO: National Assessment of Educational Progress, 52p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1977). *Mathematics techical report: Exercise volume.* (ED 138 468) (NAEP-04-MA-20) Denver, CO: National Assessment of Educational Progress, 1,412p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1976). *Mathematics technical report: Summary volume.* (ED 129 636) (NAEP-04-MA-21) Denver, CO: National Assessment of Educational Progress, 181p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1975). *Math fundamentals: Selected results from the first National Assessment of mathematics.* (ED 102 029) (NAEP-R-04-MA-01) Denver, CO: National Assessment of Educational Progress, 56p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1975). *Consumer math: Selected results from the first National Assessment of mathematics.* (ED 111 696) (NAEP-04-MA-02) Denver, CO: National Assessment of Educational Progress, 49p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1975). *National Assessment of Educational Progress. The first National Assessment of mathematics: An overview.* (ED 127 198) (NAEP-04-MA-00) Denver, CO: National Assessment of Educational Progress, 64p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Educational Testing Service. (1986). *Math objectives, 1985-86 assessment.* (17-M-10) Princeton, NJ: National Assessment of Educational Progress. 26p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Kahl, S. (1979). *Mathematical knowledge and skills: Selected results from the second assessment of mathematics.* (ED 176 964) (NAEP-09-MA-02) Denver, CO: National Assessment of Educational Progress, 83p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402 and from NAEP, CN6710, Princeton, NJ, 08541-6710.

Kahl, S. (1979). *Mathematical understanding: Selected results from the second assessment of mathematics.* (ED 182 174) (NAEP-09-MA-04) Denver, CO: National Assessment of Educational Progress, 53p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402 and from NAEP, CN6710, Princeton, NJ, 08541-6710.

Phillips, D. (1980). *Mathematics technical report: Summary volume.* (ED 186 279) (NAEP-09-MA-21) Denver, CO: National Assessment of Educational Progress, 158p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402 and from NAEP, CN6710, Princeton, NJ, 08541-6710.

Norris, E. L., & Bowes, J. E. (1970). *National Assessment of Educational Progress. mathematics objectives.* (ED 063 140) Ann Arbor, MI: National Assessment of Educational Progress, 41p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Ward, B. (1980). *Procedural handbook: 1977-78 mathematics assessment.* (ED 186 280) (NAEP-09-MA-40) Denver, CO: National Assessment of Educational Progress, 74p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710 and Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Ward, B. (1979). *Changes in mathematical achievement, 1973-78: Results from the second assessment of mathematics.* (ED 177 011) (NAEP-09-MA-01) Denver, CO: National Assessment of Educational Progress, 43p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Ward, B. (1979). *Mathematical applications: Selected results from the second assessment of mathematics.* (ED 176 965) (NAEP-09-MA-03) Denver, CO: National Assessment of Educational Progress, 66p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402 and from NAEP, CN6710, Princeton, NJ, 08541-6710.

READING

DeCrow, R. (1972). *Adult reading abilities: Definitions and measurements.* (ED 068 810) Washington, DC: National Reading Center Foundation, 6p.

Gadway, C. J. (1973). *Critical reading: Theme 8, reading.* (ED 078 387) (NAEP-02-R-08) Denver, Co: National Assessment of Educational Progress, 182p Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Gadway, C. J. (1973). *Drawing inferences: Theme 7, reading.* (ED 083 533) (NAEP-02-R-07) Denver, CO: National Assessment of Educational Progress, 275p. Available from Superintendent of Documents, U.S. Government Printing Office, Washington, DC, 20402.

Gadway, C. J. (1973). *Gleaning significant facts from passages: Theme 5, reading.* (ED 078 360) (NAEP-02-R-05) Denver, CO: National Assessment of Educational Progress, 225p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Gadway, C. J. (1973). *Main ideas and organization: Theme 6, reading.* (ED 079 688) (NAEP-02-R-06) Denver, CO: National Assessment of Educational Progress, 152p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Gadway, C. J. (1973). *Reading: Released exercises.* (ED 079 684) (NAEP-02-R-20) Denver, CO: National Assessment of Educational Progress, 424p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Gadway, C. J. (1973). *Reference materials: Theme 4, reading.* (ED 079 687) (NAEP-02-R-04) Denver, CO: National Assessment of Educational Progress, 164p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Gadway, C. J. (1973). *Understanding words and word relationships: Theme 1 of the National Assessment of reading.* (ED 077 021) (NAEP-02-R01) Denver, CO: National Assessment of Educational Progress, 106p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Gadway, C. J. (1973). *Written directions: Theme 3, reading.* (ED 083 572) (NAEP-02-R-03) Denver, CO: National Assessment of Educational Progress, 155p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Gallo, D. R. (1972). *Reading rate and comprehension: 1970-71 assessment.* (ED 076 934) (NAEP-02-R-09) Denver, CO: National Assessment of Educational Progress, 24p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1984). *Reading objectives, 1983-84 Assessment.* (NAEP 15-RL-10) (ISBN-0-88685-022-3) Denver, CO: National Assessment of Educational Progress, 11p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1982). *Reading comprehension of American youth: Do they understand what they read? Results from the 1979-80 National Assessment of reading and literature.* (ED 217 396) (NAEP-11-R-02) (ISBN-0-89398-223-7) Denver, CO: National Assessment of Educational Progress, 89p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1981). *Has Title I improved education for disadvantaged students? Evidence from three national assessments of reading.* (ED 201 995) (NAEP-SY-DS-50) Denver, CO: National Assessment of Educational Progress, 9p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1981). *Procedural handbook: 1979-80 reading and literature assessment.* (ED 210 300) (NAEP-11-RL-40) (ISBN-0-89398-221-0) Denver, CO: National Assessment of Educational Progress, 103p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1981). *Reading/literature released exercise set, 1979-80 Assessment.* (ED 205 588) (NAEP-11-RL-25) (ISBN-0-89398-222-9) Denver, CO: National Assessment of Educational Progress, 355p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1981). *Reading/literature released exercise set, 1979-80 Assessment. Supplement.* (ED 205 589) (NAEP-11-RL-26) (ISBN-0-89398-222-9) Denver, CO: National Assessment of Educational Progress, 471p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1981). *Reading, thinking and writing: Results from the 1979-80 National Assessment of reading and literature.* (ED 209 641) (NAEP-11-L-01) (ISBN-0-89398-110-9) Denver, CO: National Assessment of Educational Progress, 82p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1981). *Three National Assessments of reading changes in performance, 1970-80.* (ED 200 898) (NAEP-11-R-01) (ISBN-0-89398-220-2) Denver, CO: National Assessment of Educational Progress, 91p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1980). *Reading and literature objectives, 1979-80 Assessment. No.11-RL-10.* (ED 185 503) (NAEP-11-RL-10) Denver, CO: National Assessment of Educational Progress, 25p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1979). *The first assessment of reading, 1970-71 assessment. Released exercise set No. 02-R-25.* (ED 191 017) Denver, CO: National Assessment of Educational Progress, 341p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1978). *Reading change, 1970-75: Summary volume. reading report No.06-R-21.* (ED 153 196) Denver, CO: National Assessment of Educational Progress, 95p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1976). *Reading in America: A perspective on two assessments.* (ED 128 785) (NAEP-06-R-01) Denver, CO: National Assessment of Educational Progress, 35p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1974). *Reading objectives; Second Assessment.* (ED 089 238) Denver, CO: National Assessment of Educational Progress, 21p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1974). *Reading: Summary data.* (ED 094 359) (NAEP-02-R-00) Denver, CO: National Assessment of Educational Progress, 62p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1973). *Graphic materials: Theme 2, reading.* (ED 078 384) (NAEP-02-R-02) Denver, CO: National Assessment of Educational Progress, 213p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Educational Testing Service. (1987). *Reading objectives, 1986 and 1988 assessments.* (NAEP 17/19-R-10) (ISBN-0-88685-063-0) Princeton, NJ: National Assessment of Educational Progress, 31p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Educational Testing Service. (1985). *The reading report card: Progress toward excellence in our schools. Trends in reading over four National Assessments, 1971-1984.* (NAEP-15-R-01) (ISBN-0-88685-034-7) Princeton, NJ: National Assessment of Educational Progress, 72p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Johnson, S. S. (1973). *A survey of reading habits: Theme 4,literature.* (ED 078 366) (NAEP-02-L-04) Denver, CO: National Assessment of Educational Progress, 157p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Kalk, J. M. (1981). *The relationship between reading related background variables and comprehension achievement: A secondary analysis of the 1979-80 reading/literature assessment data.* (ED 214 132) Denver, CO: National Assessment of Educational Progress, 159p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Norris, E. L. & Bowes, J. E. (1970). *Reading objectives.* (ED 041 010) Ann Arbor, MI: National Assessment of Educational Progress, 40p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

SCIENCE

Committee on Assessing the Progress of Education. (1969). *Science objectives.* (ED 033 872) Ann Arbor, MI: National Assessment of Educational Progress, 33p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Dochterman, C. L. (1970). *National Assessment of Educational Progress. Summary of report 1. Science: National results.* (ED 043 099) Denver, CO: Education Commission of the States, 24p.

Education Commission of the States. (1979). *Science objectives for the third assessment.* (ED 179 402) Denver, CO: National Assessment of Educational Progress, 68p.

Education Comission of the States. (1978). *Science achievement in the schools: A summary of results from the 1976-77 National Assessment of science.* (ED 164 337) (NAEP-SR-08-S-01) Denver, CO: National Assessment of Educational Progress, 64p.

Education Commission of the States. (1978). *Three National Assessments of science: Changes in achievement, 1969-77. Selected results from the third National Assessment of science.* (ED 159 026) (NAEP-08-S-00) Denver, CO: National Assessment of Educational Progress, 44p.

Education Commission of the States. (1978). *The third assessment of science, 1976-77. Release exercise set.* (ED 161 686) Denver, CO: National Assessment of Educational Progress, 350p.

Education Commission of the States. (1977). *National Assessment of Educational Progress. Science technical report: Summary volume.* (ED 141 135) (NAEP-04-S-21) Denver, CO: National Assessment of Educational Progress, 201p. Available from Superintendent of Documents, U.S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1975). *National Assessment of Educational Progress. Changes in science performance, 1969-1973: Exercise volume.* (ED 127 199) Denver, CO: National Assessment of Educational Progress, 332p. Available from Superintendent of Documents, U.S. Government Printing Office, Washington, DC, 20402.

Education Commission of the States. (1975). *National Assessment of Educational Progress. released science exercises, 1972-73.* (ED 193 040) Denver, CO: National Assessment of Educational Progress, 482p.

Education Commission of the States. (1975). *National Assessments of science, 1969 and 1973: A capsule description of changes in science achievement.* (ED 104 723) (NAEP-04-S-00) Denver, CO: National Assessment of Educational Progress, 26p. Available from Superintendent of Documents, U.S Government Printing Office, Washington, DC, 20402.

Educational Testing Service. (1986). *Science objectives, 1985-86 assessment.* (17-S-10) (ISBN-0-88685-040-1) Princeton, NJ: National Assessment of Educational Progress, 40p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Mastie, M. M., & Johnson, G. H. (1972). *National Assessment of Educational Progress, science objectives for 1972-73 assessment.* (ED 072 976) Denver, CO: National Assessment of Educational Progress, 53p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

National Assessment of Educational Progress. (1970). *National Assessment of Educational Progress, report 1--Science: National results. Observations and commentary of a panel of reviewers.* (ED 055 785) Ann Arbor, MI: National Assessment of Educational Progress, 42p. Available from Superintendent of Documents, U. S. Government Printing Office, Washington, DC, 20402.

Norris, E. (1970). *National Assessment of Educational Progress, report 1--Science: National results and illustrations of group comparisons.* (ED 055 786) Ann Arbor, MI: National Assessment of Educational Progress, 226p. Available from Superintendent of Documents, U.S. Government Printing Office, Washington, DC, 20402.

Searls, D. T., & Others. (1975). *Change in national science achievement: Some analysis problems.* (ED 157 709) Denver, CO: National Assessment of Educational Progress, 19p.

OTHER REPORTS

Applebee, A. N., Langer J. A., & Mullis, I. V. S. (1987). *Learning to be literate in America: Reading, writing, and reasoning.* (No. 15-RW-01) Princeton, NJ: National Assessment of Educational Progress, 47p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Beaton, A. E. (1987). *Implementing the new design: The 1983-84 NAEP technical report.* (ISBN 0-88685-062-2) (NAEP 15-TR-20) Princeton, NJ: National Assessment of Educational Progress, 768p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Education Commission of the States. (1982). *Introduction to the National Assessment of Educational Progress public use data tapes.* National Assessment of Educational Progress, 62p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Educational Testing Service. (1987). *Profiling American education.* (No. 18-GIY) Princeton, NJ: National Assessment of Educational Progress, 20p. Available from NAEP, Educational Testing Service, Rosedale Road, Princeton, NJ, 08541-0001.

Educational Testing Service. (1987). *Learning by doing: A manual for teaching and assessing higher-order thinking in science and mathematics.* (No. 17-HOS-80) (ISBN 0-99685-059-2) Princeton, NJ: National Assessment of Educational Progress, 32p. Available from NAEP, CN6720, Princeton, NJ 08541-6710.

Educational Testing Service. (1986). *Foundations of literacy: A description of the assessment of a basic knowledge of U. S. history and literature.* (17-HL-11) (ISBN-0-88685-041-X) Princeton, NJ: National Assessment of Educational Progress, 22p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Educational Testing Service. (1986). *A framework for assessing computer competence: Defining objectives.* (17-CC-10) (ISBN-0-88685-043-6) Princeton, NJ: National Assessment of Educational Progress, 38p. Available from NAEP, CN6710, Princeton, NJ 08541-6710.

Kirsch, I. S., & Jungeblut, A. (1986). *Literacy: Profiles of America's young adults.* (No. 16-PL-02) (ISBN 0-88685-054-1) Princeton, NJ: National Assessment of Educational Progress, 68p. Available from NAEP, Educational Testing Service, Rosedale Road, Princeton, NJ, 08541.

Kirsch, I. S., & Jungeblut, A. (1986). *Literacy: Profiles of America's young adults--Final report.* (No. 16-PL-01) (ISBN 0-88685-056-8) Princeton, NJ: National Assessment of Educational Progress. Available from NAEP, Educational Testing Service, Rosedale Road, Princeton, NJ, 08541.

Messick, S., Beaton, A., & Lord, F. (1983). *NAEP reconsidered: A new design for a new era.* (NAEP-83-1) Princeton, NJ: Educational Testing Service, 93p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.

Ward, B. (1980). *The National Assessment approach to objectives and exercise development.* (ED 194 577) (NAEP-12-IP-55) Denver, CO: National Assessment of Educational Progress, 50p. Available from NAEP, CN6710, Princeton, NJ, 08541-6710.